

DETERMINING QUASAR BLACK HOLE MASS FUNCTIONS FROM THEIR BROAD EMISSION LINES: APPLICATION TO THE BRIGHT QUASAR SURVEY

BRANDON C. KELLY^{1,2,3,4}, MARIANNE VESTERGAARD⁵, XIAOHUI FAN³

Draft version November 12, 2008

ABSTRACT

We describe a Bayesian approach to estimating quasar black hole mass functions (BHMF) when using the broad emission lines to estimate black hole mass. We show how using the broad line mass estimates in combination with statistical techniques developed for luminosity function estimation (e.g., the $1/V_a$ correction) leads to statistically biased results. We derive the likelihood function for the BHMF based on the broad line mass estimates, and derive the posterior distribution for the BHMF, given the observed data. We develop our statistical approach for a flexible model where the BHMF is modelled as a mixture of Gaussian functions. Statistical inference is performed using markov chain monte carlo (MCMC) methods, and we describe a Metropolis-Hasting algorithm to perform the MCMC. The MCMC simulates random draws from the probability distribution of the BHMF parameters, given the data, and we use a simulated data set to show how these random draws may be used to estimate the probability distribution for the BHMF. In addition, we show how the MCMC output may be used to estimate the probability distribution of any quantities derived from the BHMF, such as the peak in the space density of quasars. Our method has the advantage that it is able to constrain the BHMF even beyond the survey detection limits at the adopted confidence level, accounts for measurement errors and the intrinsic uncertainty in broad line mass estimates, and provides a natural way of estimating the probability distribution of any quantities derived from the BHMF. We conclude by using our method to estimate the local active BHMF using the $z < 0.5$ Bright Quasar Survey sources. At $z \sim 0.2$, the quasar BHMF falls off approximately as a power law with slope ~ 2 for $M_{BH} \gtrsim 10^8 M_\odot$. Our analysis implies that at a given M_{BH} , $z < 0.5$ broad line quasars have a typical Eddington ratio of ~ 0.4 and a dispersion in Eddington ratio of $\lesssim 0.5$ dex.

Subject headings: galaxies: active — galaxies: mass function — galaxies: statistics — methods: data analysis — methods: numerical — methods: statistical

1. INTRODUCTION

It is widely accepted that the extraordinary activity associated with quasars¹ involves accretion onto a supermassive black hole (SMBH). The correlation between SMBH mass and both host galaxy luminosity (e.g., Kormendy & Richstone 1995; Magorrian et al. 1998; McLure & Dunlop 2001; Marconi & Hunt 2003) and stellar velocity dispersion (M_{BH} - σ relationship, e.g., Gebhardt et al. 2000; Merritt & Ferrarese 2001; Tremaine et al. 2002), together with the fact that quasars have been observed to reside in early-type galaxies (McLure et al. 1999; Kukula et al. 2001; McLeod & McLeod 2001; Nolan et al. 2001; Percival et al. 2001; Dunlop et al. 2003), implies that the evolution of spheroidal galaxies and quasars is intricately tied together (e.g., Silk & Rees 1998; Haehnelt & Kauffmann 2000; Merritt & Poon 2004; Di Matteo et al. 2005; Hopkins et al. 2006). Therefore, investigating the evolution of active super-massive black holes (SMBHs) is an important task of modern astronomy, giving insight into the importance of AGN activity on the formation of structure in the universe. Determination of the comoving number density, energy density, and mass density of active black holes is a powerful probe of the quasar-galaxy connection and the evolution of active black holes.

Recently, advances in reverberation mapping (e.g., Peterson et al. 2004) have made it possible to estimate the masses of black holes for broad line AGN. A correlation has been found between the size of the region emitting the broad lines and the luminosity of the AGN (Kaspi et al. 2005; Bentz et al. 2006), allowing one to use the source luminosity to estimate the distance between the broad line region (BLR) and the central black hole. In addition, one can estimate the velocity dispersion of the BLR gas from the broad emission line width. One then combines the BLR size estimate with the velocity estimate to obtain a virial black hole mass as $M_{BH} \propto L^b V^2$, where $b \approx 1/2$ (e.g., Wandel et al. 1999; McLure & Jarvis 2002; Vestergaard 2002; Vestergaard & Peterson 2006). Estimates of M_{BH} obtained from the broad emission lines have been used to estimate the distribution of quasar black hole masses at a variety of redshifts (e.g., McLure & Dunlop 2004; Vestergaard 2004; Kollmeier et al. 2006; Wang et al. 2006; Greene & Ho 2007;

¹ bckelly@cfa.harvard.edu

² Hubble Fellow

³ Department of Astronomy, University of Arizona, Tucson, AZ 85721

⁴ Harvard-Smithsonian Center for Astrophysics, 60 Garden St, Cambridge, MA 02138

⁵ Department of Physics and Astronomy, Tufts University, Medford, MA 02155

¹ Throughout this work we will use the terms quasar and AGN to refer generically to broad line AGNs. No luminosity difference between the two is assumed.

Vestergaard et al. 2008; Fine et al. 2008).

Given the importance of the BHMF as an observational constraint on models of quasar evolution, it is essential that a statistically accurate approach be employed when estimating the BHMF. However, the existence of complicated selection functions hinders this. A variety of methods have been used to accurately account for the selection function when estimating the quasar luminosity function. These include various binning methods (e.g., Schmidt 1968; Avni & Bahcall 1980; Page & Carrera 2000), maximum-likelihood fitting (e.g., Marshall et al. 1983; Fan et al. 2001), a semi-parametric approach (Schafer 2007), and Bayesian approaches (e.g., Andreon et al. 2005; Kelly et al. 2008, hereafter KJV08). In addition, there have been a variety of methods proposed for estimating the cumulative distribution function of the luminosity function (e.g., Lynden-Bell 1971; Efron & Petrosian 1992; Maloney & Petrosian 1999). While these techniques have been effective for estimating luminosity functions, estimating the BHMF from the broad line mass estimates is a more difficult problem, and currently there does not exist a statistically correct method of estimating the BHMF.

If we could directly measure black hole mass for quasars, and if the selection function only depended on M_{BH} and z , then we could simply employ the formalism developed for luminosity function estimation, after replacing L with M_{BH} . However, surveys are selected based on luminosity and redshift, not on M_{BH} . At any given luminosity there exists a range in black hole mass, and thus one cannot simply employ the luminosity selection function ‘as-is’ to correct for the flux limit. In other words, completeness in flux is not the same thing as completeness in M_{BH} , and the use of a flux selection results in a softer selection function for M_{BH} . Moreover, we cannot directly observe M_{BH} for large samples of quasars, but rather derive an estimate of M_{BH} from their broad emission lines. The intrinsic uncertainty on M_{BH} derived from the broad emission lines is ~ 0.4 dex (Vestergaard & Peterson 2006), and the uncertainty on M_{BH} broadens the inferred distribution of M_{BH} (e.g., Kelly & Bechtold 2007; Shen et al. 2007; Fine et al. 2008). As a result, even if there is no flux limit, the BHMF inferred directly from the broad line mass estimates will be systematically underestimated near the peak and overestimated in the tails. In order to ensure an accurate estimate of the BHMF it is important to correct for the uncertainty in the estimates of M_{BH} .

Motivated by these issues, we have developed a Bayesian method for estimating the BHMF. In KJV08 we derived the likelihood function and posterior probability distribution for luminosity function estimation, and we described a mixture of Gaussian functions model for the luminosity function. In this work, we extend our statistical method and derive the likelihood function of the BHMF by relating the observed data to the true BHMF, and derive the posterior probability distribution of the BHMF parameters, given the observed data. While the likelihood function and posterior are valid for any parametric form, we focus on a flexible parametric model where the BHMF is modeled as a sum of Gaussian functions. This is a type of ‘non-parametric’ approach, where the basic idea is that the individual Gaussian functions do not have any physical meaning, but that given enough Gaussian functions one can obtain a suitably accurate approximation to the true BHMF. Modeling the BHMF as a mixture of normals avoids the problem of choosing a particular parametric form, especially in the absence of any guidance from astrophysical theory. In addition, we describe a markov chain monte carlo (MCMC) algorithm for obtaining random draws from the posterior distribution. These random draws allow one to estimate the posterior distribution for the BHMF, as well as any quantities derived from it. The MCMC method therefore allows a straight-forward method of calculating errors on any quantity derived from the BHMF. Because the Bayesian approach is valid for any sample size, one is able to place reliable constraints on the BHMF and related quantities, even where the survey becomes incomplete.

Because of the diversity and mathematical complexity of some parts of this paper, we summarize the main results here. We do this so that the reader who is only interested in specific aspects of this paper can conveniently consult the sections of interest.

- In § 2.2 we derive the general form of the likelihood function for black hole mass function estimation based on quasar broad emission lines. Because we can not directly observe M_{BH} for a large sample of quasars, the likelihood function gives the probability of observing a set of redshifts, luminosities, and line widths, given an assumed BHMF. In § 2.3 we derive the black hole mass selection function, and discuss how the differences between the M_{BH} selection function and the luminosity selection function affect estimating the BHMF. The reader who is interested in the likelihood function of the broad line quasar BHMF, or issues regarding correcting for incompleteness in M_{BH} , should consult this section.
- In § 3 we describe a Bayesian approach to black hole mass function estimation. We build on the likelihood function derived in § 2.2 to derive the probability distribution of the BHMF, given the observed data (i.e., the posterior distribution). The reader who is interested in a Bayesian approach to BHMF estimation should consult this section.
- In § 4 we develop a mixture of Gaussian functions model for the black hole mass function, deriving the likelihood function and posterior distribution for this model. Under this model, the BHMF is modelled as a weighted sum of Gaussian functions. This model has the advantage that, given a suitably large enough number of Gaussian functions, it is flexible enough to give an accurate estimate of any smooth and continuous BHMF. This allows the model to adapt to the true BHMF, thus minimizing the bias that can result when assuming a parametric form for the BHMF. In addition, we also describe our statistical model for the distribution of luminosities at a given M_{BH} , and the distribution of line widths at a given L and M_{BH} . These two distributions are necessary in order to link the BHMF to the observed set of luminosities and line widths. The reader who is interested in employing our mixture of Gaussian functions model should consult this section.

- Because of the large number of parameters associated with black hole mass function estimation, Bayesian inference is most easily performed by obtaining random draws of the BHMF from the posterior distribution. In § 5 we describe a Metropolis-Hastings algorithm (MHA) for obtaining random draws of the BHMF from the posterior distribution, assuming our mixture of Gaussian functions model. The reader who is interested in the computational aspects of ‘fitting’ the mixture of Gaussian functions model, or who is interested in the computational aspects of Bayesian inference for the BHMF, should consult this section.
- In § 6 we use simulation to illustrate the effectiveness of our Bayesian Gaussian mixture model for black hole mass function estimation. We construct a simulated data set similar to the Sloan Digital Sky Survey DR3 Quasar Cataloge (Schneider et al. 2005). We then use our mixture of Gaussian functions model to recover the true BHMF and show that our mixture model is able to place reliable constraints on the BHMF over all values of M_{BH} . In contrast, we show that estimating the BHMF by binning up the broad line mass estimates, and applying a simple $1/V_a$ correction, systematically biases the inferred BHMF toward larger M_{BH} . We also illustrate how to use the MHA output to constrain any quantity derived from the BHMF, and how to use the MHA output to assess the quality of the fit. Finally, we discuss difficulties associated with inferring the distribution of Eddington ratios. The reader who is interested in assessing the effectiveness of our statistical approach, or who is interested in using the MHA output for statistical inference on the BHMF, should consult this section.
- In § 7 we use our statistical method to estimate the $z < 0.5$ BHMF from the Bright Quasar Survey sources. We also attempt to infer the mean and dispersion in the $z < 0.5$ distribution of Eddington ratios. The reader who is interested in the scientific results regarding our estimated $z < 0.5$ BHMF should consult this section.

We adopt a cosmology based on the the WMAP best-fit parameters ($h = 0.71, \Omega_m = 0.27, \Omega_\Lambda = 0.73$, Spergel et al. 2003)

2. THE LIKELIHOOD FUNCTION

2.1. NOTATION

We use the common statistical notation that an estimate of a quantity is denoted by placing a ‘hat’ above it; e.g., $\hat{\theta}$ is an estimate of the true value of the parameter θ . We denote a normal distribution with mean μ and variance σ^2 as $N(\mu, \sigma^2)$, and we denote as $N_p(\mu, \Sigma)$ a multivariate normal distribution with p -element mean vector μ and $p \times p$ covariance matrix Σ . If we want to explicitly identify the argument of the Gaussian function, we use the notation $N(x|\mu, \sigma^2)$, which should be understood to be a Gaussian function with mean μ and variance σ^2 as a function of x . We will often use the common statistical notation where “ \sim ” means “is drawn from” or “is distributed as”. This should not be confused with the common usage of “ \sim ” implying “similar to”. For example, $x \sim N(\mu, \sigma^2)$ states that x is drawn from a normal distribution with mean μ and variance σ^2 , whereas $x \sim 1$ states that the value of x is similar to one.

2.2. Likelihood Function for the BHMF Estimated from AGN Broad Emission Lines

The black hole mass function, denoted as $\phi(M_{BH}, z)dM_{BH}$, is the number of sources per comoving volume $V(z)$ with black hole masses in the range $M_{BH}, M_{BH} + dM_{BH}$. The black hole mass function is related to the probability distribution of (M_{BH}, z) by

$$p(M_{BH}, z) = \frac{1}{N} \phi(M_{BH}, z) \frac{dV}{dz}, \quad (1)$$

where N is the total number of sources in the universe, and is given by the integral of ϕ over M_{BH} and $V(z)$. If we assume a parametric form for $\phi(M_{BH}, z)$, with parameters θ , we can derive the likelihood function for the observed data. The likelihood function is the probability of observing one’s data, given the assumed model. The presense of selection effects and intrinsic uncertainty in the broad line mass estimates can make this difficult, as the observed data likelihood function is not simply given by Equation (1). However, we can account for these difficulties by first deriving the likelihood function for the complete set of data, and then integrating over the missing data to obtain the observed data likelihood function.

For broad line AGNs, we can relate the distribution of M_{BH} and z to the joint distribution of L_λ, \mathbf{v} , and z . Here, $\mathbf{v} = (v_{H\beta}, v_{MgII}, v_{CIV})$, where $v_{H\beta} = v_{H\beta}$ is the the velocity dispersion for the H β broad line emitting gas, and similarly for v_{MgII} and v_{CIV} . These three lines are commonly used in estimating M_{BH} from single-epoch spectra of broad line AGN (e.g., McLure & Jarvis 2002; Kaspi et al. 2005; Vestergaard 2002; Vestergaard & Peterson 2006), where the velocity dispersion is typically estimated from the *FWHM* of the emission line. The distribution of L_λ and \mathbf{v} are then related to the BHMF via the *R-L* relationship and the virial theorem.

The BHMF for broad line AGN can be inferred from the distribution of L_λ, \mathbf{v} , and z , and thus it is necessary to formulate the observed data likelihood function in terms of $(L_\lambda, \mathbf{v}, z)$. While it is possible to formulate the likelihood function in terms of the broad line mass estimates, denoted as $\hat{M}_{BL} \propto L_\lambda^{1/2} V^2$, the logarithm of the broad line mass estimates are simply linear combinations of $\log L_\lambda$ and $\log \mathbf{v}$, and thus statistical inference does not depend on whether we formulate the likelihood function in terms of L_λ and \mathbf{v} or \hat{M}_{BL} . We find it mathematically simpler and more intuitive to infer the BHMF directly from the distribution of L_λ, \mathbf{v} , and z , as opposed to inferring it from the distribution of L_λ, \hat{M}_{BL} , and z .

Following the discussion in KJV08, we derive the likelihood function for the set of observed luminosities, redshifts, and emission line widths. We introduce an indicator variable I denoting whether a source is included in the survey or not: if $I_i = 1$ then a source is included, otherwise, $I_i = 0$. The variable I is considered to be part of the observed data in the sense that we ‘observe’ whether a source is detected or not. The survey selection function is the probability of including the i^{th} source in one’s survey, $p(I_i = 1|\mathbf{v}_i, L_{\lambda,i}, z_i)$. Here, we have assumed that the probability of including a source in one’s sample only depends on luminosity, redshift, and emission line width, and is therefore conditionally independent of M_{BH} . This is the case, in general, since one can only select a survey based on quantities that are directly observable. Including the additional ‘data’ I , the observed data likelihood function for broad line AGN is:

$$p(\mathbf{v}_{obs}, L_{obs}, z_{obs}, I|\theta, N) \propto \quad (2)$$

$$C_n^N \prod_{i \in \mathcal{A}_{obs}} \int p(\mathbf{v}_i, L_{\lambda,i}, M_{BH,i}, z_i|\theta) dM_{BH,i} \quad (3)$$

$$\times \prod_{j \in \mathcal{A}_{mis}} \int \int \int \int p(I = 0|\mathbf{v}_j, L_{\lambda,j}, z_j) p(\mathbf{v}_j, L_{\lambda,j}, M_{BH,j}, z_j|\theta) d\mathbf{v}_j dL_{\lambda,j} dM_{BH,j} dz_j \quad (4)$$

$$\propto C_n^N [p(I = 0|\theta)]^{N-n} \prod_{i \in \mathcal{A}_{obs}} p(\mathbf{v}_i, L_{\lambda,i}, z_i|\theta), \quad (5)$$

where \mathcal{A}_{obs} denotes the set of sources included in one’s survey, \mathcal{A}_{mis} denotes the set of sources not included in one’s survey, and on the last line we have omitted terms that do not depend on N or θ . Here,

$$p(\mathbf{v}_i, L_{\lambda,i}, z_i|\theta) = \int_0^\infty p(\mathbf{v}_i, L_{\lambda,i}, z_i, M_{BH,i}|\theta) dM_{BH,i} \quad (6)$$

is the probability of observing values of $\mathbf{v}_i, L_{\lambda,i}$, and z_i for the i^{th} source, given θ , and

$$p(I = 0|\theta) = \int_0^\infty \int_0^\infty \int_0^\infty p(I = 0|\mathbf{v}, L_\lambda, z) p(\mathbf{v}, L_\lambda, z|\theta) d\mathbf{v} dL_\lambda dz \quad (7)$$

is the probability that the survey misses a source, given θ ; note that $p(I = 0|\theta) = 1 - p(I = 1|\theta)$. Qualitatively, the observed data likelihood function for the BHMF is the probability of observing a set of n emission line widths $\mathbf{v}_1, \dots, \mathbf{v}_n$, luminosities $L_{\lambda,1}, \dots, L_{\lambda,n}$, and redshifts z_1, \dots, z_n given the assumed BHMF model parameterized by θ , multiplied by the probability of not detecting $N - n$ sources given θ , multiplied by the number of ways to select a subset of n sources from a set of N total sources. Equation (5) can be maximized to calculate a maximum likelihood estimate of the black hole mass function when using broad line estimates of M_{BH} , or combined with a prior distribution to perform Bayesian inference.

It is often preferred to write the BHMF observed data likelihood function by factoring the joint distribution of $\mathbf{v}, L_\lambda, M_{BH}$, and z into conditional distributions. This has the advantage of being easier to interpret and work with, especially when attempting to connect the distribution of line widths and luminosities to the distribution of black hole mass. The joint distribution can be factored as (Kelly & Bechtold 2007)

$$p(\mathbf{v}, L_\lambda, M_{BH}, z) = p(\mathbf{v}|L_\lambda, M_{BH}, z) p(L_\lambda|M_{BH}, z) p(M_{BH}, z). \quad (8)$$

Here, $p(\mathbf{v}|L_\lambda, M_{BH}, z)$ is the distribution of emission line widths at a given L_λ, M_{BH} , and z , $p(L_\lambda|M_{BH}, z)$ is the distribution of luminosities at a given M_{BH} and z , and $p(M_{BH}, z)$ is the probability distribution of black hole mass and redshift, related to the BHMF via Equation (1). When using broad line estimates of M_{BH} , it is assumed that $p(\mathbf{v}|L_\lambda, M_{BH}, z)$ is set by the virial theorem, where the distance between the central black hole and the broad line-emitting gas depends on L_λ via the R - L relationship. In this work we assume that the R - L relationship does not depend on z (e.g., Vestergaard 2004), and thus $p(\mathbf{v}|L_\lambda, M_{BH}, z) = p(\mathbf{v}|L_\lambda, M_{BH})$.

Under the factorization given by Equation (8), the observed data likelihood function (Eq. [5]) becomes

$$p(\mathbf{v}_{obs}, L_{obs}, z_{obs}, I|\theta, N) \propto C_n^N [p(I = 0|\theta)]^{N-n} \prod_{i \in \mathcal{A}_{obs}} \int_0^\infty p(\mathbf{v}_i|L_{\lambda,i}, M_{BH,i}, \theta) p(L_{\lambda,i}|M_{BH,i}, z_i, \theta) p(M_{BH,i}, z_i|\theta) dM_{BH,i}. \quad (9)$$

The BHMF likelihood function, given by Equation (5) or (9), is entirely general, and it is necessary to assume parametric forms in order to make use of it. In § 4 we describe a parametric form based on a mixture of Gaussian functions model, and explicitly calculate Equation (5) for the mixture model.

2.3. Selection Function

The selection probability, $p(I = 1|\mathbf{v}, L_\lambda, z)$, depends on both the luminosity and redshift through the usual flux dependence, but can also depend on the emission line width. In particular, an upper limit on \mathbf{v} may occur if there is a width above which emission lines become difficult to distinguish from the continuum and iron emission. In this case, if all emission lines in one’s spectrum are wider than the maximum line width than one is not able to obtain a reliable estimate of the line width for any emission line, and therefore the source is not used to estimate $\phi(M_{BH}, z)$. A lower

limit on the line width may be imposed in order to prevent the inclusion of narrow line AGN, for which broad line mass estimates are not valid. In this case the inclusion criterion might be that at least one emission line is broader than, say, $FWHM = 2000 \text{ km s}^{-1}$. In addition to the limits on line width that may be imposed, there is an upper and lower limit on z due to redshifting of emission lines out of the observable spectral range. For example, if one uses optical spectra then the range of useable spectra is $0 < z \lesssim 4.5$, as the C IV line redshifts into the near-infrared for $z \gtrsim 4.5$.

Denote the upper and lower limit of \mathbf{v} as v_{min} and v_{max} , and the upper and lower limit of z as z_{min} and z_{max} . Furthermore, denote the usual survey selection function in terms of L_λ and z as $s(L_\lambda, z)$, where $s(L_\lambda, z)$ is the probability that a source is included in the survey before any cuts on line width are imposed; $s(L_\lambda, z)$ would typically correspond to the selection function used in luminosity function estimation. Note that in this work $s(L_\lambda, z)$ gives the probability that any source in the universe is included in the survey, given its luminosity and redshift, and thus $s(L_\lambda, z) \leq \Omega/4\pi$, where $\Omega/4\pi$ is the fraction of the sky covered by the survey. Then, $p(I_i = 1|\mathbf{v}_i, L_{\lambda,i}, z_i) = s(L_{\lambda,i}, z_i)$ if $z_{min} \leq z_i \leq z_{max}$ and at least one emission line has $v_{min} \leq v_i \leq v_{max}$; otherwise, $p(I_i = 1|\mathbf{v}_i, L_{\lambda,i}, z_i) = 0$. In this case, the probability that a source is included in the survey (see Eq.[7]) is

$$p(I = 1|\theta) = \int_0^\infty \int_{z_{min}}^{z_{max}} s(L_\lambda, z) \int_{v_{min}}^{v_{max}} \int_0^\infty p(\mathbf{v}|L_\lambda, M_{BH}, \theta) p(L_\lambda|M_{BH}, z, \theta) p(M_{BH}, z|\theta) dM_{BH} d\mathbf{v} dz dL_\lambda, \quad (10)$$

where the inner two integrals are over \mathbf{v} and M_{BH} , and the outer two integrals are over L_λ and z . One can then insert Equation (10) into Equation (5) to get the likelihood function.

It is informative to express the selection function in terms of black hole mass and redshift. The selection function as a function of black hole mass and redshift is the probability of including a source, given its M_{BH} and z , and is calculated as

$$p(I = 1|M_{BH}, z) = \int_0^\infty s(L_\lambda, z) p(L_\lambda|M_{BH}, z) \int_{v_{min}}^{v_{max}} p(\mathbf{v}|L_\lambda, M_{BH}) dL_\lambda d\mathbf{v}. \quad (11)$$

At any given value of M_{BH} a range of luminosities and emission line widths are possible, and thus sources with low black hole mass can be detected if they are bright enough and have line widths $v_{min} < v < v_{max}$. Conversely, sources with high black hole masses can be missed by the survey if their luminosity is below the flux limit at that redshift, or if their line width falls outside of the detectable range. This has the effect of smoothing the survey's selection function, and thus the black hole mass selection function is a broadened form of the flux selection function.

As an example, consider the case when the selection function is simply a flux limit. In this case, the selection function is

$$s(l, z) = \begin{cases} 1 & \text{if } 4\pi f_{min} D_L^2(z) < L_\lambda < 4\pi f_{max} D_L^2(z) \\ 0 & \text{otherwise} \end{cases}, \quad (12)$$

where f_{min} is the survey's lower flux limit, f_{max} is the survey's upper flux limit, and $D_L(z)$ is the luminosity distance to redshift z . For simplicity, in this example we assume that there is no additional cut on emission line width, i.e., $v_{min} = 0$ and $v_{max} = \infty$. In this case, the black hole mass selection function, $p(I = 1|M_{BH}, z)$, is the convolution of the luminosity selection function with the distribution of L_λ at a given M_{BH} . If the distribution of $\log L_\lambda$ at a given M_{BH} is a Gaussian function with mean $\alpha_0 + \alpha_m \log M_{BH}$ and dispersion σ_l , then the black hole mass selection function is

$$p(I = 1|M_{BH}, z) = \Phi\left(\frac{\log L_{max}(z) - \alpha_0 - \alpha_m \log M_{BH}}{\sigma_l}\right) - \Phi\left(\frac{\log L_{min}(z) - \alpha_0 - \alpha_m \log M_{BH}}{\sigma_l}\right). \quad (13)$$

Here, $L_{max}(z) = 4\pi f_{max} D_L^2(z)$, $L_{min}(z) = 4\pi f_{min} D_L^2(z)$, and $\Phi(\cdot)$ is the cumulative distribution function of the standard normal distribution.

In Figure 1 we show the black hole mass selection function, $p(I = 1|M_{BH}, z)$, given by Equation (13) at $z = 1$. Here, we have used the SDSS quasar sample flux limit, $19.1 > i > 15$, $\alpha_0 = 37$, $\alpha_m = 1$, and $\sigma_l = 0.6$ dex. Because the black hole mass selection function is the convolution of the luminosity selection function with the distribution of L_λ at a given M_{BH} , the black hole mass selection function is positive over a wider range in M_{BH} , as compared to the range in L_λ for which $s(L_\lambda, z)$ is positive. However, because $p(I = 1|M_{BH}, z)$ spreads the selection probability over a wider range in M_{BH} , all bins in M_{BH} are incomplete.

The difference in selection functions for black hole mass and luminosity results in an important distinction between the estimation of black hole mass functions and the estimation of luminosity functions. First, one cannot correct the binned BHMF for the survey flux limits by simply applying the $1/V_a$ correction. This is a common technique used for estimating binned luminosity functions, where the number density in a (L_λ, z) bin is corrected using the survey volume in which a source with luminosity L_λ could have been detected and still remained in the redshift bin. In the case of the BHMF, a survey volume in which the black hole could have been detected ceases to have any meaning, as black holes can be detected over many different survey volumes, albeit with varying probability. Alternatively, the $1/V_a$ correction can be thought of as dividing the number of sources in a bin in (L_λ, z) by the detection probability as a function of L_λ and z . Therefore, applying a $1/V_a$ correction to a bin in (M_{BH}, z) is essentially the same as dividing the number of sources in a bin in (M_{BH}, z) by the detection probability as a function of L_λ and z . For the simple example shown in Figure 1, those quasars in a given bin in (M_{BH}, z) that happen to have luminosities $L_{min}(z) < L_\lambda < L_{max}(z)$ will

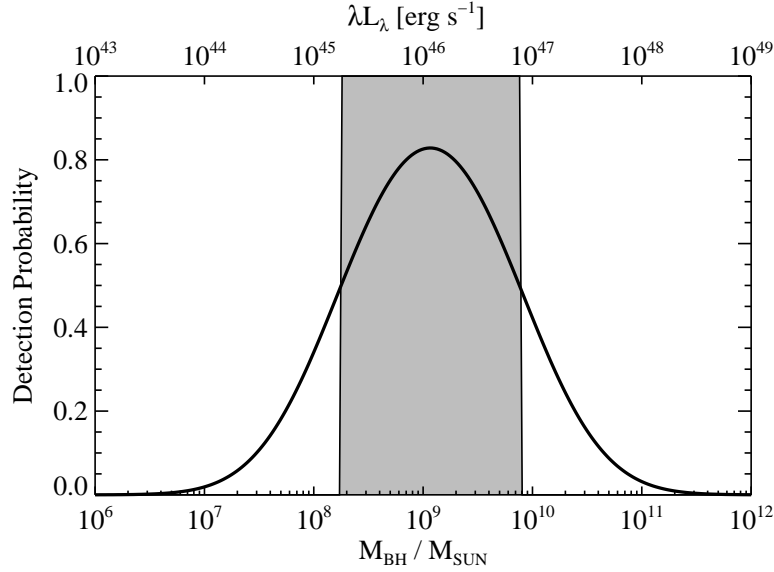


FIG. 1.— Comparison of the selection function for luminosity (shaded region) and black hole mass (curve) for a simple upper and lower flux limit. The selection is complete in luminosity within the flux limits, but is never ‘complete’ in M_{BH} . The intrinsic physical range in luminosity at a given black hole mass creates a more complicated selection function for M_{BH} , since at any given M_{BH} and z only those quasars with luminosities within the flux limits are detected.

receive no correction, since $s(L_\lambda, z) = 1$. However, those quasars which have luminosities outside of the detectable range will not be detected. The end result is a systematic underestimate of the binned BHMF.

The number of sources in a given bin in the M_{BH} - z plane can be estimated by dividing the observed number of black holes in each bin by the black hole mass selection function, $p(I = 1|M_{BH}, z)$. Similarly, one can use a $1/V_a$ -type correction by calculating an ‘effective’ $1/V_a$, found by integrating dV/dz over the black hole mass selection function. This approach has been adopted previously within the context of binned luminosity functions (e.g., Warren et al. 1994; Fan et al. 2001). However, it is essential that the black hole mass selection function be used and *not* the luminosity selection function. Unfortunately, this implies that one must assume a form for $p(L_\lambda|M_{BH}, z)$.

3. POSTERIOR DISTRIBUTION FOR THE BHMF PARAMETERS

The posterior probability distribution of the model parameters is

$$p(\theta, N|\mathbf{v}_{obs}, L_{obs}, z_{obs}, I) \propto p(\theta, N)p(\mathbf{v}_{obs}, L_{obs}, z_{obs}, I|\theta, N), \quad (14)$$

where $p(\theta, N)$ is the prior on (θ, N) , and $p(\mathbf{v}_{obs}, L_{obs}, z_{obs}, I|\theta, N)$ is the likelihood function, given by Equation (5). The posterior distribution gives the probability that θ and N have a given value, given the observed data $(\mathbf{v}_{obs}, L_{obs}, z_{obs})$. Therefore, the posterior distribution of θ and N can be used to obtain the probability that $\phi(M_{BH}, z)$ has any given value, given that we have observed some set of emission line widths, luminosities, and redshifts.

It is of use to decompose the posterior as $p(N, \theta|x_{obs}) \propto p(N|\theta, x_{obs})p(\theta|x_{obs})$, where we have abbreviated the observed data as $x_{obs} = (\mathbf{v}_{obs}, L_{obs}, z_{obs})$. This decomposition separates the posterior into the conditional posterior of the BHMF normalization, $p(N|\theta, x_{obs})$, from the marginal posterior of the BHMF shape, $p(\theta|x_{obs})$. In this work we take N and θ to be independent in their prior distribution, $p(N, \theta) = p(N)p(\theta)$, and that the prior on N is uniform over $\log N$. In this case, one can show (e.g., Gelman et al. 2004, KJV08) that the marginal posterior distribution of θ is

$$p(\theta|\mathbf{v}_{obs}, L_{obs}, z_{obs}) \propto p(\theta) [p(I = 1|\theta)]^{-n} \prod_{i \in \mathcal{A}_{obs}} p(\mathbf{v}_i, L_{\lambda, i}, z_i|\theta), \quad (15)$$

where $p(I = 1|\theta) = 1 - p(I = 0|\theta)$.

Under the prior $p(\log N) \propto 1$, the conditional posterior of $N|\theta, x_{obs}$ is a negative binomial distribution with parameters n and $p(I = 1|\theta)$. The negative binomial distribution gives the probability that the total number of sources is equal to N , given that there have been n detections with probability of detection $p(I = 1|\theta)$:

$$p(N|n, \theta) = C_{n-1}^{N-1} [p(I = 1|\theta)]^n [p(I = 0|\theta)]^{N-n}. \quad (16)$$

Because of the large number of parameters in the model, Bayesian inference is most easily performed by randomly drawing values of N and θ from their posterior. Based on the decomposition $p(\theta, N|x_{obs}) \propto p(N|n, \theta)p(\theta|x_{obs})$, we can obtain random draws of (N, θ) by first drawing values of θ from Equation (15). Then, for each draw of θ , we draw a value of N from the negative binomial distribution. Random draws for θ may be obtained via markov chain monte carlo (MCMC) methods, described in § 5, and random draws from the negative binomial distribution are easily obtained using standard methods (e.g., Gelman et al. 2004, KJV08).

4. THE STATISTICAL MODEL

In order to compute the likelihood function for the observed set of luminosities, redshifts, and broad emission line widths (see Eq.[9]), it is necessary to relate the BHMF to the distribution of L_λ and \mathbf{v} . To do this, Equation (8) implies that we need three terms. The first term is an assumed BHMF, $p(M_{BH}, z) = N^{-1}(dV/dz)^{-1}\phi(M_{BH}, z)$. The second term is an assumed distribution of luminosities at a given black hole mass and redshift, $p(L_\lambda|M_{BH}, z)$. The third term is an assumed distribution of broad emission line widths at a given luminosity and black hole mass, $p(\mathbf{v}|L_\lambda, M_{BH})$. Once we have a parameteric form for each of these three distributions, we can calculate the observed data likelihood directly from Equation (9). In this section we describe parameteric forms for each of these distributions based on a mixture of Gaussian functions model.

4.1. Mixture of Gaussian Functions Model for the BHMF

The mixture of Gaussian functions model is a common ‘non-parametric’ model that allows flexibility when estimating the BHMF. The basic idea is that one can use a suitably large enough number of Gaussian functions to accurately approximate the true BHMF, even though the individual Gaussian functions have no physical meaning. Furthermore, the Gaussian mixture model is also conjugate to the distributions $p(L_\lambda|m)$ and $p(\mathbf{v}|L_\lambda, m)$ assumed in §§4.2 and 4.3, thus enabling us to calculate some of the integrals in Equation (15) analytically.

In KJV08 we described a mixture of Gaussian functions model for a luminosity function. The mixture of Gaussian functions model of the BHMF is identical to that for the luminosity function, after replacing L with M_{BH} . Our mixture of Gaussian functions model, including our adopted prior, is described in KJV08; for completeness we briefly review it here.

The mixture of K Gaussian functions model for the BHMF is

$$p(\log M_{BH}, \log z|\pi, \mu, \Sigma) = \sum_{k=1}^K \frac{\pi_k}{2\pi|\Sigma_k|^{1/2}} \exp \left[-\frac{1}{2}(\mathbf{y} - \mu_k)^T \Sigma_k^{-1}(\mathbf{y} - \mu_k) \right], \quad (17)$$

where $\sum_{k=1}^K \pi_k = 1$. Here, $\mathbf{y} = (\log M_{BH}, \log z)$, μ_k is the 2-element mean vector for the k^{th} Gaussian functions, Σ_k is the 2×2 covariance matrix for the k^{th} Gaussian function, and x^T denotes the transpose of x . In addition, we denote $\pi = (\pi_1, \dots, \pi_K)$, $\mu = (\mu_1, \dots, \mu_K)$, and $\Sigma = (\Sigma_1, \dots, \Sigma_K)$. The variance in $\log M_{BH}$ for Gaussian function k is $\sigma_{m,k}^2 = \Sigma_{11,k}$, the variance in $\log z$ for Gaussian function k is $\sigma_{z,k}^2 = \Sigma_{22,k}$, and the covariance between $\log M_{BH}$ and $\log z$ for Gaussian function k is $\sigma_{mz,k} = \Sigma_{12,k}$. Note that Equation(17) is equivalent to assuming that $p(M_{BH}, z)$ is a mixture of log-normal densities. Under the mixture model, the BHMF can be calculated from Equations (1) and (17). Noting that $p(M_{BH}, z) = p(\log M_{BH}, \log z)/(M_{BH}z(\ln 10)^2)$, the mixture of normals model for the BHMF is

$$\phi(M_{BH}, z|\theta, N) = \frac{N}{M_{BH}z(\ln 10)^2} \left(\frac{dV}{dz} \right)^{-1} \sum_{k=1}^K \frac{\pi_k}{2\pi|\Sigma_k|^{1/2}} \exp \left[-\frac{1}{2}(\mathbf{y} - \mu_k)^T \Sigma_k^{-1}(\mathbf{y} - \mu_k) \right], \quad (18)$$

where, as before, $\mathbf{y} = (\log M_{BH}, \log z)$.

4.2. The Distribution of L_λ at a Given M_{BH}

We model the distribution of luminosities at a given M_{BH} as a log-normal distribution, where the average $\log L_\lambda$ at a given M_{BH} depends linearly on $\log M_{BH}$:

$$p(\log L_\lambda|M_{BH}, \alpha) = \frac{1}{\sqrt{2\pi\sigma_l^2}} \exp \left[-\frac{1}{2} \left(\frac{\log L_\lambda - \alpha_0 - \alpha_m \log M_{BH}}{\sigma_l} \right)^2 \right]. \quad (19)$$

Here, the unknown parameters are $\alpha = (\alpha_0, \alpha_m, \sigma_l^2)$. This is equivalent to assuming a simple linear regression of $\log L_\lambda$ on $\log M_{BH}$, where α_0 is the constant, α_m is the slope, and σ_l is the standard deviation of the random Gaussian dispersion about the regression line. We assume a uniform prior on these parameters, i.e., $p(\alpha_0, \alpha_m, \sigma_l) \propto 1$.

The form of the M_{BH} - L_λ relationship given by Equation (19) is motivated by noting that L_λ can be related to M_{BH} as

$$\lambda L_\lambda = 1.3 \times 10^{38} \frac{\Gamma_{Edd}}{C_\lambda} \frac{M_{BH}}{M_\odot} \text{ [erg s}^{-1}\text{]}, \quad (20)$$

where $\Gamma_{Edd} \equiv L_{bol}/L_{Edd}$ is the Eddington ratio, and C_λ is the bolometric correction to λL_λ . Equation (20) implies that the distribution of luminosities at a given black hole mass is caused by the distribution in Eddington ratios and bolometric corrections at a given black hole mass. The distribution of $\log L_\lambda$ at a given M_{BH} is the convolution of the distribution of $\log \Gamma_{Edd}$ at a given M_{BH} , with the distribution of $\log C_\lambda$ at a given M_{BH} . The parameter σ_l is thus an estimate of the dispersion in $\log(\Gamma_{Edd}/C_\lambda)$ at a given M_{BH} .

If both Γ_{Edd} and C_λ are statistically independent of M_{BH} , then we would expect that on average $L_\lambda \propto M_{BH}$, i.e., $\alpha_m = 1$. However, if Γ_{Edd} or C_λ are correlated with M_{BH} , then $\alpha_m \neq 1$. Currently, it is unknown whether M_{BH} and Γ_{Edd} are correlated. However, it is likely that quasar SEDs depend on both Γ_{Edd} and M_{BH} , and therefore the bolometric correction will also depend on Γ_{Edd} and M_{BH} . Indeed, recently some authors have found evidence that the bolometric correction depends on Eddington ratio (Vasudevan & Fabian 2007) and black hole mass (Kelly et al. 2008).

Therefore, it is likely that $\alpha_m \neq 1$, and we leave it as a free parameter. In addition, comparison of Equation (19) with Equation (20), and assuming that Γ_{Edd}/C_λ is independent of M_{BH} , implies that the average value of Γ_{Edd}/C_λ is related to α_0 according to $E(\log \Gamma_{Edd}/C_\lambda) = \alpha_0 - 38.11$, where $E(x)$ denotes the expectation value of x . Therefore, one can use α_0 to estimate the typical broad line quasar Eddington ratio, assuming a typical bolometric correction.

Currently, there is little known about the distribution of luminosities at a given black hole mass, so for simplicity we assume the simple linear form given by Equation (19). Furthermore, the assumption of a Gaussian distribution in $\log L$ at a given M_{BH} is consistent with the L - M_{BH} relationship for those AGN with reverberation mapping data (Kelly & Bechtold 2007). More sophisticated models could include a non-linear dependence on $\log M_{BH}$, an additional redshift dependence, or non-Gaussian distribution. Unfortunately, this introduces additional complexity into the model. Furthermore, an additional redshift dependence in Equation (19) implies that the distribution of Γ_{Edd} or C_λ at a given M_{BH} evolves. However, currently most investigations have not found any evidence for significant evolution in Γ_{Edd} (e.g., Corbett et al. 2003; McLure & Dunlop 2004; Vestergaard 2004; Kollmeier et al. 2006), and it is unclear if the quasar SED evolves at a given M_{BH} . Therefore, there is currently no compelling evidence to justify inclusion of a redshift dependence in Equation (19). In addition, we note that it is impossible to use $p(L|M_{BH})$ to infer the distribution of Eddington ratios without making an assumption about the distribution of C_{bol} , as Equation (20) shows that Γ_{Edd} and C_{bol} are degenerate. While estimating the distribution of Γ_{Edd} is of significant interest, it is beyond the scope of this work to develop a robust technique to do so, as our goal is to estimate the black hole mass function.

Because of the large number of parameters, large uncertainty in the broad line black hole mass estimates, and flux limit, estimating the BHMF is already a difficult statistical problem. As such, our approach is to initially assume the simple form given by Equation (19) in order to keep the degrees of freedom low, and to check if this assumption is consistent with our data (see § 6.3). If it is found that the observed data are inconsistent with this statistical model (e.g., see § 6.3) then Equation (19) should be modified.

4.3. The Distribution of v at a given L and M_{BH}

Following Kelly & Bechtold (2007), we can derive the distribution of emission line widths at a given luminosity and black hole mass. Given an AGN luminosity, L_λ^{BL} , the BLR distance R is assumed to be set by the luminosity according to the R - L relationship, $R \propto L_\lambda^{\beta_l}$, with some additional log-normal statistical scatter:

$$p(\log R | L_\lambda^{BL}) = \frac{1}{\sqrt{2\pi\sigma_r^2}} \exp \left[-\frac{1}{2} \left(\frac{\log R - r_0 - \beta_l \log L_\lambda^{BL}}{\sigma_r} \right)^2 \right]. \quad (21)$$

Here, r_0 is a constant, σ_r is the dispersion in $\log R$ at a given luminosity, and L_λ^{BL} is the AGN continuum luminosity at some reference wavelength appropriate for the broad emission line of interest. Note that the reference wavelength for L_λ^{BL} is not necessarily the same wavelength as for L_λ used in § 4.2. In particular, the wavelength for L_λ used in the M_{BH} - L_λ relationship should be chosen to adequately account for the selection function, while the reference wavelength for L_λ^{BL} should be appropriate for describing the R - L relationship. Since AGN continua are well described by a power-law, $f_\nu \propto \nu^{-\alpha}$, it should be easy to calculate L_λ at different values so long as the spectral index, α , is known. The intrinsic scatter in R at a given L_λ^{BL} is likely due to variations in quasar SED, reddening, non-instantaneous response of the BLR to continuum variations, etc.

Assuming that the BLR gas is gravitationally bound, the velocity dispersion of the broad line-emitting gas is related to R and M_{BH} as $M_{BH} = f R v^2 / G$. Here, G is the gravitational constant, and f is a factor that converts the virial product, $R M_{BH} / G$, to a mass. We do not directly measure v , but instead estimate it by the $FWHM$ or dispersion of the broad emission line in a single-epoch spectra. As a result, the measured line width will scatter about the actual value of v , where this scatter may be due in part to variations in line profile shape and the existence of stationary components in the single-epoch line profile. In our statistical model we assume that this scatter is log-normal with a dispersion of σ_v . In addition, the value of f depends on the measure of line width used. Onken et al. (2004) estimated f by comparing black hole masses derived from reverberation mapping with those derived from the M_{BH} - σ relationship, and find that on average $f = 1.4 \pm 0.4$ when using the $FWHM$. This value is consistent with a value of $f = 0.75$ expected from a spherical BLR geometry (e.g., Netzer 1990).

Under our model, the distribution of emission line widths at a given BLR size and black hole mass is

$$\log v | R, M_{BH} = \frac{1}{\sqrt{2\pi\sigma_v^2}} \exp \left\{ -\frac{1}{2} \left[\frac{\log v - v_0 - 1/2(\log f + \log R - \log M_{BH})}{\sigma_v} \right]^2 \right\}. \quad (22)$$

where v_0 is a constant. For convenience, here and throughout this paper we denote the estimate of the BLR gas velocity dispersion as v , i.e., v is either the $FWHM$ or dispersion of the emission line. The term v in Equation (22) should not be confused with the actual velocity dispersion of the BLR gas, but is an estimate of it based on a measure of the width of the broad emission line. From Equation (22) it is apparent that the term f shifts the distribution of $\log v$ by a constant amount, which has the effect of shifting the inferred BHMF by a constant amount in $\log M_{BH}$. Throughout the rest of this work we assume the value of $f = 1.4$ found by Onken et al. (2004).

The distribution of v at a given L and M_{BH} is obtained from Equations (21) and (22) by averaging the distribution of v at a given R and M_{BH} over the distribution of R at a given L_λ^{BL} :

$$p(\log v | L_\lambda^{BL}, M_{BH}, \beta) = \int_{-\infty}^{\infty} p(\log v | \log R, M_{BH}, \beta) p(\log R | L_\lambda^{BL}, \beta) d \log R \quad (23)$$

$$= \frac{1}{\sqrt{2\pi\sigma_{BL}^2}} \exp \left\{ -\frac{1}{2} \left[\frac{\log v - \beta_0 - 1/2(\beta_l \log L_\lambda^{BL} - \log M_{BH})}{\sigma_{BL}} \right]^2 \right\}. \quad (24)$$

where β_0 is a constant, $\sigma_{BL}^2 = \sigma_v^2 + \sigma_r^2/4$, and $\beta \equiv (\beta_0, \beta_l, \sigma_{BL})$. Note that in Equation (24) we have absorbed $\log f$ into the constant term, β_0 . The term σ_{BL} is the dispersion in emission line widths at a given luminosity and black hole mass, and can be related to the intrinsic uncertainty in the broad line estimates of M_{BH} . The usual broad line mass estimates of AGN can be obtained by reexpressing the mean of Equation (24) in terms of M_{BH} : $\log \hat{M}_{BL} = \beta_l \log L_\lambda^{BL} + 2 \log v - 2\beta_0$, or equivalently $\hat{M}_{BL} \propto L_{\lambda,BL}^{\beta_l} v^2$. The intrinsic uncertainty on the broad line mass estimates is set by a combination of the intrinsic dispersion in R and at a given L , and the uncertainty in using the single-epoch line width as an estimate of the broad line gas velocity dispersion: $\sigma_{\hat{M}_{BL}} = 2\sigma_{BL}$. Equation (24) describes the statistical uncertainty in the broad line mass estimates, and does not account for any additional systematic errors (e.g., Krolik 2001; Collin et al. 2006).

It is typically the case that one employs multiple emission lines to estimate M_{BH} , producing black hole mass estimates across a broad range of redshifts and luminosities. In our work, we use the H β , Mg II, and C IV emission lines. In order to facilitate the use of different emission lines in the BHM estimation, we introduce an indicator variable denoted by δ . Here, $\delta_{H\beta} = 1$ if the H β line width is available, and $\delta_{H\beta} = 0$ if the H β line widths is not available; δ_{MgII} and δ_{CIV} are defined in an equivalent manner. For example, if one is using optical spectra, then at $z = 0.4$ only the H β emission line is available, and therefore $\delta_{H\beta} = 1$, $\delta_{MgII} = 0$, and $\delta_{CIV} = 0$.

Assuming that the line width distributions for each line are independent at a given luminosity and black hole mass, then the observed distribution of line widths is the product of Equation (24) for each individual emission line:

$$p(\log \mathbf{v} | L, M_{BH}, z, \beta) = [N(\log v_{H\beta} | \bar{v}_{H\beta}, \sigma_{H\beta}^2)]^{\delta_{H\beta}} [N(\log v_{MgII} | \bar{v}_{MgII}, \sigma_{MgII}^2)]^{\delta_{MgII}} [N(\log v_{CIV} | \bar{v}_{CIV}, \sigma_{CIV}^2)]^{\delta_{CIV}} \quad (25)$$

Here, the average line width for H β is $\bar{v}_{H\beta} = \beta_0^{H\beta} - (1/2)\beta_l^{H\beta} \log L_\lambda^{H\beta} + (1/2) \log M_{BH}$, and likewise for Mg II and C IV. Here, $L_\lambda^{H\beta}$ denotes the value of L_λ that is used to calibrate the broad line mass estimates for H β , typically $L_\lambda(5100\text{\AA})$.

Vestergaard & Peterson (2006) give equations for calculating broad line mass estimates from H β and C IV, derived from the most recent reverberation mapping data (Peterson et al. 2004; Kaspi et al. 2005), and Vestergaard et al. (2008, in progress) give an equation for calculating a broad line mass estimate from Mg II. These mass scaling relationships are:

$$\log \hat{M}_{H\beta} = -21.09 + 0.50 \log \lambda L_\lambda(5100\text{\AA}) + 2 \log FWHM_{H\beta} \quad (26)$$

$$\log \hat{M}_{MgII} = -21.21 + 0.50 \log \lambda L_\lambda(2100\text{\AA}) + 2 \log FWHM_{MgII} \quad (27)$$

$$\log \hat{M}_{CIV} = -22.66 + 0.53 \log \lambda L_\lambda(1350\text{\AA}) + 2 \log FWHM_{CIV} \quad (28)$$

For the equations listed above we have used the $FWHM$ of the emission line as an estimate of the velocity dispersion, i.e., $v = FWHM$. Because $\log \hat{M}_{BL} = \beta_l \log \lambda L_\lambda^{BL} + 2 \log v - 2\beta_0$, it follows that $\beta_0^{H\beta} = 10.55$, $\beta_0^{MgII} = 10.61$, $\beta_0^{CIV} = 11.33$, and $\beta_l \approx 0.5$ for all three emission lines. In addition, Vestergaard & Peterson (2006) find the statistical uncertainty in the broad line mass estimates to be 0.43 dex and 0.36 dex for H β and C IV, respectively. Therefore, since $\sigma_{BL} = \sigma_{\hat{M}_{BL}}/2$, it follows that $\sigma_{H\beta} \approx 0.22$ and $\sigma_{CIV} \approx 0.18$ dex. Likewise, Vestergaard et al. (2008, in progress) find the intrinsic uncertainty in the broad line mass estimate for Mg II to be ~ 0.4 dex, and therefore $\sigma_{MgII} \approx 0.2$ dex. However, this statistical uncertainty may be smaller if a correction is made in the virial relationship for radiation pressure (Marconi et al. 2008).

Broad line mass estimates are now fairly well understood, and we derive our prior distribution for β from the scaling results of Vestergaard & Peterson (2006) and Vestergaard et al. (2008, in progress). We fix $\beta_l = 0.5, 0.5$, and 0.53 for H β , Mg II, and C IV, respectively. However, in order to account for the uncertainty in these scaling relationships, we consider β_0 and σ_{BL} to be free parameters in our model. We cannot estimate the normalization and statistical uncertainty in the broad line mass estimates solely from the distribution of \mathbf{v} , L , and z , since β_0 and σ_{BL} are degenerate with the other parameters. Therefore, it is necessary to place constraints on β_0 and σ_{BL} through a prior distribution. This allows us to constrain β_0 and σ_{BL} while still incorporating their uncertainty. The parameters for the prior distribution of β_0 and σ_{BL} are based on the uncertainty in the scaling relationships of Vestergaard & Peterson (2006) and Vestergaard et al. (2008, in progress). Our prior for β_0 are independent Gaussian distributions with means equal to 10.55, 10.61, and 11.33 for H β , Mg II, and C IV, respectively, and standard deviations equal to 0.1. To allow greater flexibility in our model, we chose the prior standard deviation of 0.1 to be wider than the formal uncertainty on the scaling factors of ≈ 0.02 reported by Vestergaard & Peterson (2006). For each emission line, our prior for σ_{BL} is a scaled inverse- χ^2 distribution with $\nu = 25$ degrees of freedom and scale parameter equal to 0.2 dex. We chose $\nu = 25$ degrees of freedom because approximately 25 AGN were used to derive the scaling relationships in Vestergaard & Peterson (2006). The values of β_0 were constrained to be within ± 0.3 (i.e., $\pm 3\sigma$) of the values reported by Vestergaard & Peterson (2006) and Vestergaard et al. (2008, in progress), and the values of σ_{BL} were constrained to be within the interval containing 99% of the probability for the scaled inverse- χ^2 distribution. By placing these constraints on β_0 and σ_{BL} , we ensure that their values remain consistent with the results derived from reverberation mapping.

4.4. Likelihood function for Mixture of Gaussian Functions Model

Now that we have formulated the conditional distributions, we can calculate the likelihood function for the mixture of Gaussian functions model of $\phi(M_{BH}, z)$. Comparison with Equation (15) suggests that we need two terms: $p(\mathbf{v}_i, L_{\lambda,i}, z_i|\theta)$ and $p(I = 1|\theta)$. The first term is the joint distribution of line widths, luminosities, and redshifts:

$$p(\mathbf{v}_i, L_{\lambda,i}, z_i|\theta) = \int p(\mathbf{v}_i|L_{\lambda,i}, M_{BH,i}, \beta) p(L_{\lambda,i}|M_{BH,i}, \alpha) p(M_{BH,i}, z_i|\pi, \mu, \Sigma) dM_{BH,i}, \quad (29)$$

where $\theta = (\alpha, \beta, \pi, \mu, \Sigma)$.

The integral in Equation (29) can be done analytically by inserting Equations (17), (19), and (25) into Equation (29). However, the result depends on the number of emission lines used for the i^{th} source. Expressing the likelihood function for a single emission line in terms of logarithms, $p(\log v_i, \log L_{\lambda,i}, \log z_i|\theta)$ is a mixture of K 3-dimensional Gaussian functions:

$$p(\log v_i, \log L_{\lambda,i}, \log z_i|\theta) = \sum_{k=1}^K \frac{\pi_k}{\sqrt{8\pi^3|V_k|}} \exp \left\{ -\frac{1}{2}(\mathbf{x}_i - \xi_k)^T V_k^{-1}(\mathbf{x}_i - \xi_k) \right\} \quad (30)$$

$$\mathbf{x}_i = (\log v_i, \log L_{\lambda,i}, \log z_i) \quad (31)$$

$$\xi_k = (\bar{v}_k, \bar{l}_k, \mu_{z,k}) \quad (32)$$

$$\bar{l}_k = \alpha_0 + \alpha_m \mu_{m,k} \quad (33)$$

$$\bar{v}_k = \beta_0 - \frac{1}{2}\beta_l \bar{l}_{BL,k} + \frac{1}{2}\mu_{m,k} \quad (34)$$

$$\bar{l}_{BL,k} = \bar{l}_k + (1 + \alpha_\lambda) \log \left(\frac{\lambda_{BL}}{\lambda_{ML}} \right) \quad (35)$$

$$V_k = \begin{pmatrix} \text{Var}(\log v|k) & \text{Cov}(\log v, \log l|k) & \text{Cov}(\log v, \log z|k) \\ \text{Cov}(\log v, \log l|k) & \text{Var}(\log l|k) & \alpha_m \sigma_{mz,k} \\ \text{Cov}(\log v, \log z|k) & \alpha_m \sigma_{mz,k} & \sigma_{z,k}^2 \end{pmatrix} \quad (36)$$

$$\text{Var}(\log v|k) = \sigma_{BL}^2 + \frac{1}{4}[\beta_l^2 \text{Var}(\log l|k) + (1 - \alpha_m)\sigma_{m,k}^2] \quad (37)$$

$$\text{Var}(\log l|k) = \sigma_l^2 + \alpha_m^2 \sigma_{m,k}^2 \quad (38)$$

$$\text{Cov}(\log v, \log l|k) = \frac{1}{2}\alpha_m \sigma_{m,k}^2 - \frac{1}{2}\beta_l \text{Var}(\log l|k) \quad (39)$$

$$\text{Cov}(\log v, \log z|k) = \left(\frac{1}{2} - \frac{1}{2}\beta_l \alpha_m \right) \sigma_{mz,k}. \quad (40)$$

Here, ξ_k and V_k are the mean vector and covariance matrix of $(\log v_i, \log L_{\lambda,i}, \log z_i)$ for the k^{th} Gaussian function, respectively. In addition, \bar{l}_k is the mean $\log L_\lambda$ for Gaussian function k , \bar{v}_k is the mean v for Gaussian function k , \bar{l}_{BL} is the mean $\log L_\lambda^{BL}$ for Gaussian function k , $\text{Var}(\log v|k)$ is the variance in $\log v$ for Gaussian function k , $\text{Var}(\log L_\lambda|k)$ is the variance in $\log L_\lambda$ for Gaussian function k , $\text{Cov}(\log v, \log L_\lambda|k)$ is the covariance between $\log v$ and $\log L_\lambda$ for Gaussian function k , and $\text{Cov}(\log v, \log z|k)$ is the covariance between $\log v$ and $\log z$ for Gaussian function k ; note that $\alpha_m \sigma_{mz,k}$ is the covariance between $\log L_\lambda$ and z for Gaussian function k . The mean $\log L_\lambda^{BL}$ for Gaussian function k is calculated from \bar{l}_k assuming a power-law continuum of the form $L_\lambda^{BL} = L_\lambda (\lambda_{BL}/\lambda_{ML})^{\alpha_\lambda}$, where λ_{BL} is the wavelength used in the $R-L_\lambda^{BL}$ relationship for the emission line of interest, and λ_{ML} is the wavelength that the $M_{BH}-L_\lambda$ is formulated in. For example, $\lambda_{BL} = 5100\text{\AA}$ for the H β -based mass scaling relationship of Vestergaard & Peterson (2006), and λ_{ML} may be, say, equal to 2500\AA . Note that we are assuming that α_λ is known.

In Equation (30) it should be understood that v_i, β_0, β_l , and σ_{BL}^2 correspond to the particular emission line being used. For example, if one is using the C IV line width for the i^{th} source, then $v_i = v_{CIV,i}$, $\beta_0 = \beta_0^{CIV}$, $\beta_l = \beta_l^{CIV}$, and $\sigma_{BL}^2 = \sigma_{CIV}^2$.

If there are two emission line widths available for the i^{th} AGN, then $p(\mathbf{v}_i, L_{\lambda,i}, z_i|\theta)$ is a mixture of K 4-dimensional Gaussian functions:

$$p(\log \mathbf{v}_i, \log L_{\lambda,i}, z_i|\theta) = \sum_{k=1}^K \frac{\pi_k}{\sqrt{16\pi^4|V_k|}} \exp \left\{ -\frac{1}{2}(\mathbf{x}_i - \xi_k)^T V_k^{-1}(\mathbf{x}_i - \xi_k) \right\} \quad (41)$$

$$\mathbf{x}_i = (\log v_{1,i}, \log v_{2,i}, \log L_{\lambda,i}, \log z_i) \quad (42)$$

$$\xi_k = (\bar{v}_{1,k}, \bar{v}_{2,k}, \bar{l}_k, \mu_{z,k}) \quad (43)$$

$$\text{Cov}(\log v_1, \log v_2|k) = \frac{1}{4}(\beta_{l,1}\beta_{l,2}\text{Var}(\log L_\lambda|k) + \sigma_m^2) \quad (44)$$

Here, $\text{Cov}(\log v_1, \log v_2|k)$ denotes the covariance between the logarithms of the two line widths, v_1 and v_2 , for the k^{th}

Gaussian function. The 4×4 covariance matrix of $(\log \mathbf{v}_i, \log L_{\lambda,i}, \log z_i)$ is

$$V_k = \begin{pmatrix} \text{Var}(\log v_1|k) & \text{Cov}(\log v_1, \log v_2|k) & \text{Cov}(\log v_1, \log L_\lambda|k) & \text{Cov}(\log v_1, \log z|k) \\ \text{Cov}(\log v_1, \log v_2|k) & \text{Var}(\log v_2|k) & \text{Cov}(\log v_2, \log L_\lambda|k) & \text{Cov}(\log v_2, \log z|k) \\ \text{Cov}(\log v_1, \log L_\lambda|k) & \text{Cov}(\log v_2, \log L_\lambda|k) & \text{Var}(\log L_\lambda|k) & \alpha_m \sigma_{mz,k} \\ \text{Cov}(\log v_1, \log z|k) & \text{Cov}(\log v_2, \log z|k) & \alpha_m \sigma_{mz,k} & \sigma_{z,k}^2 \end{pmatrix}. \quad (45)$$

The other terms are given by Equations (34)–(40), where it should be understood that β_0, β_l , and σ_{BL}^2 correspond to the values appropriate for each emission line. For example, at $z \sim 0.6$ both H β and Mg II are observable in the optical spectral region, and thus it is possible to have line widths for both emission lines. In this case, $v_{1,i}$ is the logarithm of the H β width for the i^{th} source, $v_{2,i}$ is the logarithm of the Mg II width for the i^{th} source, $\beta_{l,1}$ corresponds to β_l for the H β line, and $\beta_{l,2}$ corresponds to β_l for the Mg II line. The labeling of the H β line width as v_1 is irrelevant, and the same result would be obtained if we had labeled the H β line width as v_2 .

It should be noted that in Equation (41) we have made the assumption that if *at least* one emission line has $v_{\min} < v < v_{\max}$, then v is estimated for all emission lines in the observable spectral range at that redshift. If this is not the case, then Equation (41) must be integrated over $v_{1,i}$ or $v_{2,i}$ if either of $v_{1,i}$ or $v_{2,i}$ fall outside of (v_{\min}, v_{\max}) .

The term $p(I = 1|\theta)$ is the probability that a source is included in one's sample for a given set of model parameters θ . Under the mixture of Gaussian functions model, Equation (10) can be simplified, allowing more efficient calculation. However, as above, the actual functional form of $p(I = 0|\theta)$ depends on the number of emission lines used in broad line mass estimation. If only one emission line is used, then Equation (10) becomes

$$p(I = 1|\theta) = \int_{-\infty}^{\infty} \int_{z_{\min}}^{z_{\max}} \frac{s(L_\lambda, z)}{z \ln 10} \sum_{k=1}^K \pi_k f_{\mathbf{v}}(L_\lambda, z, \theta, k) N_2(\mathbf{y}_{lz} | \bar{\mathbf{y}}_{lz,k}, V_{lz,k}) dz dL_\lambda \quad (46)$$

$$\mathbf{y}_{lz} = (\log L_\lambda, \log z) \quad (47)$$

$$\bar{\mathbf{y}}_{lz,k} = (\bar{l}_k, \mu_{z,k}) \quad (48)$$

$$V_{lz,k} = \begin{pmatrix} \text{Var}(\log L_\lambda|k) & \alpha_m \sigma_{mz,k} \\ \alpha_m \sigma_{mz,k} & \sigma_{z,k}^2 \end{pmatrix}. \quad (49)$$

The term $f_{\mathbf{v}}(L_\lambda, z, \theta, k)$ is the probability that a source has at least one line width between v_{\min} and v_{\max} for the k^{th} Gaussian function, given its luminosity and redshift. For redshifts where only one emission line is used, $f_{\mathbf{v}}(L_\lambda, z, \theta, k) = \text{Pr}(v_{\min} < v < v_{\max} | L_\lambda, z, \theta, k)$, where

$$\text{Pr}(v_{\min} < v < v_{\max} | L_\lambda, z, \theta, k) = \Phi \left(\frac{\log v_{\max} - E(\log v | L_\lambda, z, k)}{\sqrt{\text{Var}(\log v | L_\lambda, z, k)}} \right) - \Phi \left(\frac{\log v_{\min} - E(\log v | L_\lambda, z, k)}{\sqrt{\text{Var}(\log v | L_\lambda, z, k)}} \right) \quad (50)$$

$$E(\log v | L_\lambda, z, k) = \bar{v}_k + \mathbf{c}_k^T V_{lz,k}^{-1} (\mathbf{y}_{lz} - \bar{\mathbf{y}}_{lz,k}) \quad (51)$$

$$\text{Var}(\log v | L_\lambda, z, k) = \text{Var}(\log v | k) - \mathbf{c}_k^T V_{lz,k}^{-1} \mathbf{c}_k^T \quad (52)$$

$$\mathbf{c}_k = [\text{Cov}(\log v, \log L_\lambda | k), \text{Cov}(\log v, \log z | k)]. \quad (53)$$

Here, $\Phi(\cdot)$ is the cumulative distribution function of the standard normal distribution, $E(\log v | L_\lambda, z, k)$ is the mean of $\log v$ for the k^{th} Gaussian function at a given L_λ and z , $\text{Var}(\log v | L_\lambda, z, k)$ is the variance in $\log v$ for the k^{th} Gaussian function at a given L_λ and z , and \mathbf{c}_k is a 2-dimensional vector containing the covariances between $\log v$ and both $\log L_\lambda$ and $\log z$. The standard normal cumulative distribution function can be efficiently computed using a look-up table, and therefore only two integrals need to be calculated numerically in Equation (46).

If one is using multiple emission lines for estimating $\phi(M_{BH}, z)$, then $f_{\mathbf{v}}(L_\lambda, z, \theta, k)$ must be modified to account for this. Equation (50) gives the probability that an emission line has a line width $v_{\min} < v < v_{\max}$, under the assumption that only one emission line is used at any given redshift. However, if there are redshifts where two emission lines are used, then $f_{\mathbf{v}}(L_\lambda, z, \theta, k)$ must be modified, as in these cases we need the probability that *at least* one emission line has $v_{\min} < v < v_{\max}$. At redshifts where two emission lines are used, $f_{\mathbf{v}}(L_\lambda, z, \theta, k)$ becomes the probability that either $v_{\min} < v_1 < v_{\max}$ or $v_{\min} < v_2 < v_{\max}$:

$$f_{\mathbf{v}}(L_\lambda, z, \theta, k) = \text{Pr}(v_{\min} < v_1 < v_{\max} | L_\lambda, z, \theta, k) + \text{Pr}(v_{\min} < v_2 < v_{\max} | L_\lambda, z, \theta, k) \quad (54)$$

$$- \text{Pr}(v_{\min} < v_1 < v_{\max} | L_\lambda, z, \theta, k) \text{Pr}(v_{\min} < v_2 < v_{\max} | L_\lambda, z, \theta, k), \quad (55)$$

where $\text{Pr}(v_{\min} < v_j < v_{\max} | L_\lambda, z, \theta, k)$ are given by Equation (50) for $j = 1, 2$, respectively.

As an example, at $z \sim 0.2$ only the H β line is available in the optical spectral region, and thus, at this redshift, an optical survey can only employ the H β line for estimating the BHMF. In this case, $p(\log v_i, \log L_{\lambda,i}, \log z_i | \theta)$ is given by Equation (30), and $f_{\mathbf{v}}(L_\lambda, z, \theta, k)$ is given by Equation (50). However, at $z \sim 0.6$, both H β and Mg II are observable in the optical spectral region, and thus both may be employed for estimating the BHMF. At this redshift, $p(\log \mathbf{v}_i, \log L_{\lambda,i}, \log z_i | \theta)$ is given by Equation (41), and $f_{\mathbf{v}}(L_\lambda, z, \theta, k)$ is given by Equation (55), where $\mathbf{v} = (v_1, v_2)$, v_1 is the H β line width, and v_2 is the Mg II line width (or vice versa). If only one emission line is available at any particular redshift, either because of limited spectral range or because of a choice on the part of the researcher to ignore certain emission lines, then only Equations (30) and (50) need be used.

The functional forms of $p(\mathbf{v}_i, L_{\lambda,i}, z_i|\theta)$ and $p(I=1|\theta)$ given above can be inserted into Equation (5) to obtain the likelihood function for the mixture of normals model. A maximum-likelihood estimate of $\phi(M_{BH}, z)$ can be obtained by first maximizing Equation (5) with respect to N and $\theta = (\alpha_0, \alpha_m, \sigma_l^2, \beta_0, \beta_l, \sigma_{BL}^2, \pi, \mu, \Sigma)$. Then, using the maximum-likelihood estimate of (N, π, μ, Σ) , the maximum-likelihood estimate of $\phi(M_{BH}, z)$ is calculated by using Equation (17) in Equation (1). Unfortunately, for $K > 1$ Gaussian functions, maximizing the likelihood for the Gaussian mixture model is a notoriously difficult optimization problem. The maximization is probably most efficiently performed using the Expectation-Maximization (EM, Dempster, Laird, & Rubin 1977) algorithm, or employing a stochastic search routine. Since we focus on Bayesian inference, a derivation of the EM algorithm for the BHMF is beyond the scope of this work.

The posterior distribution of θ and N can be calculated using the forms given above for $p(\log \mathbf{v}_i, \log L_{\lambda,i}, z_i|\theta)$ and $p(I=1|\theta)$. In this case, one inserts the equations for $p(\log \mathbf{v}_i, \log L_{\lambda,i}, \log z_i|\theta)$ and $p(I=1|\theta)$ for the Gaussian mixture model into Equations (15) and (16). The prior distribution, $p(\theta)$, is given by Equation (21) in KJV08.

4.5. Accounting for Measurement Error

The preceding discussion has assumed that \mathbf{v}_i and $L_{\lambda,i}$ are known. However, in general, both quantities are measured with error. The effect of measurement error is to artificially broaden the observed distributions of \mathbf{v}_i and $L_{\lambda,i}$. Because the Bayesian approach attempts to define the set of BHMFs that are consistent with the observed distribution of $\mathbf{v}_i, L_{\lambda,i}$, and z_i , where ‘consistency’ is measured by the posterior probability distribution, measurement error can affect statistical inference on the BHMF. If the variance of the measurement errors on \mathbf{v}_i and $L_{\lambda,i}$ are small compared to the intrinsic physical variance in these quantities, then measurement error does not have a significant effect on the results. In general, the measurement errors on $L_{\lambda,i}$ will likely be small compared to the physical range in AGN luminosities, so we neglect them. This may not always be the case for the emission line widths, and in this section we modify the likelihood function for the mixture of Gaussian functions model to include measurement errors in \mathbf{v}_i . The general method of handling measurement errors within a Bayesian or likelihood function approach is described in many references (e.g., Kelly 2007). For the sake of brevity, we omit the derivations and simply report the modifications to the likelihood function.

If one is only employing one emission line at a given redshift, then Equation (29) can be factored as

$$\begin{aligned} p(\log v_i, \log L_{\lambda,i}, \log z_i|\theta) &= p(\log v_i|L_{\lambda,i}, z_i, \theta)p(\log L_{\lambda,i}, \log z_i|\theta) \\ &= \sum_{k=1}^K \pi_k p(\log v_i|L_{\lambda,i}, z_i, \theta, k)p(\log L_{\lambda,i}, \log z_i|\theta, k) \end{aligned} \quad (56)$$

Under the mixture of Gaussian functions model, the joint distribution of luminosity and redshift for the k^{th} Gaussian function is obtained from Equations (30)–(36) by simply omitting the terms that depend on v_i :

$$p(\log L_{\lambda,i}, \log z_i|\theta, k) = \frac{1}{\sqrt{4\pi^2|V_{lz,k}|}} \exp \left\{ -\frac{1}{2}(\mathbf{y}_{lz,i} - \bar{\mathbf{y}}_{lz,k})^T V_{lz,k}^{-1}(\mathbf{y}_{lz,i} - \bar{\mathbf{y}}_{lz,k}) \right\}. \quad (57)$$

Here, $\mathbf{y}_{lz,i} = (\log L_{\lambda,i}, \log z_i)$, $\bar{\mathbf{y}}_{lz,k}$ is given by Equation (48) and $V_{lz,k}$ is given by Equation (49). The distribution of the measured $\log v_i$ at $L_{\lambda,i}$ and z_i for the k^{th} Gaussian function is

$$p(\log v_i|L_{\lambda,i}, z_i, \theta, k) = \frac{1}{\sqrt{2\pi[Var(\log v|L_{\lambda,i}, z_i, k) + \sigma_{v,i}^2]}} \exp \left\{ -\frac{1}{2} \frac{(\log v_i - E(\log v|L_{\lambda,i}, z_i, k))^2}{Var(\log v|L_{\lambda,i}, z_i, k) + \sigma_{v,i}^2} \right\}. \quad (58)$$

Here, $\sigma_{v,i}^2$ is the variance of the measurement error on v_i , $E(\log v|L_{\lambda,i}, z_i, k)$ is given by Equation (51), and $Var(\log v|L_{\lambda,i}, z_i, k)$ is given by Equation (52). From Equation (58) the effect of measurement error on the line width becomes apparent: the distribution of line widths at a given luminosity and redshift is broadened by an amount dependent on the magnitude of the line width measurement error. If $\sigma_{v,i}^2 \ll Var(\log v|L_{\lambda,i}, z_i, k)$ then Equation (56) reduces to Equation (29). Otherwise, if measurement error on v_i is a concern, Equations (56)–(58) should be used for Equation (29) instead of Equation (30).

If one is employing two emission lines at a given redshift, then Equation (29) becomes

$$p(\log \mathbf{v}_i, \log L_{\lambda,i}, \log z_i|\theta) = p(\log v_{1,i}|L_{\lambda,i}, z_i, \theta)p(\log v_{2,i}|L_{\lambda,i}, z_i, \theta)p(\log L_{\lambda,i}, \log z_i|\theta). \quad (59)$$

In this case, $p(\log v_{j,i}|L_{\lambda,i}, z_i, \theta)$, $j = 1, 2$, must be calculated separately for each emission line from Equation (58).

5. POSTERIOR DISTRIBUTION OF THE BHMF VIA MARKOV CHAIN MONTE CARLO

The number of free parameters in our statistical model is $6K + 8$, where K is the number of Gaussian functions used to approximate $\phi(\log M_{BH}, \log z)$. Because of the large number of free parameters, summarizing the posterior is most efficiently done by using Markov Chain Monte Carlo techniques to simulate random draws of θ and N from the posterior distribution. In this work we use the Metropolis-Hastings algorithm (MHA, Metropolis & Ulam 1949; Metropolis et al. 1953; Hastings 1970) to perform the MCMC. We use the MHA to obtain a set of random draws from the marginal posterior distribution of θ , given by Equation (15). Then, given the values of θ , random draws for N

may be obtained from the negative binomial distribution. A further description of the Metropolis-Hastings algorithm is given by KJV08, and our MHA is an extension of the MHA described in KJV08. For further details on the MHA see Chib & Greenberg (1995) or Gelman et al. (2004).

As in KJV08, we denote the current value of a parameter by placing a \sim over its symbol, and we denote the proposal value by placing a $*$ in the superscript. For example, if one were updating α_0 , then $\tilde{\alpha}_0$ denotes the current value of α_0 in the random walk, α_0^* denotes the proposed value of α_0 , $\tilde{\theta}$ denotes the current value of θ , and θ^* denotes the proposed value of θ , i.e., $\theta^* = (\alpha_0^*, \tilde{\alpha}_m, \tilde{\sigma}_l^2, \tilde{\beta}_0, \tilde{\sigma}_{BL}^2, \tilde{\pi}, \tilde{\mu}, \tilde{\Sigma}, \tilde{\mu}_0, \tilde{A}, \tilde{T})$. Here, μ_0 , A and T are the parameters for the prior distribution on the mixture of Gaussian functions parameter (see KJV08). In addition, for ease of notation we define $x_{obs} = (\mathbf{v}_{obs}, L_{obs}, z_{obs})$ to be the set of observable quantities.

Our adopted Metropolis-Hastings algorithm is as follows:

1. Start with initial guesses for $\alpha_0, \alpha_m, \sigma_l^2, \beta_0, \sigma_{BL}^2, \pi, \mu, \Sigma, \mu_0$, and A .
2. Draw a proposal value for α_0 and α_m from a 2-dimensional normal distribution centered at the current values of α_0 and α_m with set covariance matrix, Σ_α . The proposal values of α_0 and α_m are then simulated as $(\alpha_0^*, \alpha_m^*) \sim N_2([\tilde{\alpha}_0, \tilde{\alpha}_m], \Sigma_\alpha)$. If $p(\theta^*|x_{obs}) > p(\tilde{\theta}|x_{obs})$ then set $\tilde{\alpha}_0 = \alpha_0^*$ and $\tilde{\alpha}_m = \alpha_m^*$. Otherwise, calculate the ratio $r_\alpha = p(\theta^*|x_{obs})/p(\tilde{\theta}|x_{obs})$ and draw a random number uniformly distributed between 0 and 1, denoted as u . If $u < r_\alpha$ then set $\tilde{\alpha}_0 = \alpha_0^*$ and $\tilde{\alpha}_m = \alpha_m^*$. Otherwise, if $u > r_\alpha$, the values of $\tilde{\alpha}_0$ and $\tilde{\alpha}_m$ remain unchanged.
3. Draw a proposal value for $\log \sigma_l^2$ as $\log \tilde{\sigma}_l^2 \sim N(2 \log \sigma_l^*, \sigma_{\sigma_l}^2)$, where $\sigma_{\sigma_l}^2$ is some set variance. Similar to before, calculate the ratio $r_\sigma = \sigma_l^* p(\theta^*|x_{obs}) / \tilde{\sigma}_l p(\tilde{\theta}|x_{obs})$. Here, the term $\sigma_l^* / \tilde{\sigma}_l$ arises because the MHA acceptance rule must be corrected for the asymmetry in the log-normal jumping distribution used for σ_l^2 . If $r_\sigma \geq 1$ then set $\tilde{\sigma}_l = \sigma_l^*$, otherwise set $\tilde{\sigma}_l = \sigma_l^*$ with probability r_σ . This is done by drawing a uniformly distributed random variable as in step 2.
4. Draw a proposal value for β_0 from a normal distribution centered at the current value of β_0 with set variance, σ_β^2 . If $p(\theta^*|x_{obs}) > p(\tilde{\theta}|x_{obs})$ then set $\tilde{\beta}_0 = \beta_0^*$. Otherwise, calculate the ratio $r_\beta = p(\theta^*|x_{obs})/p(\tilde{\theta}|x_{obs})$ and draw a random number uniformly distributed between 0 and 1, denoted as u . If $u < r_\beta$ then set $\tilde{\beta}_0 = \beta_0^*$. Otherwise, if $u > r_\beta$, then the value of $\tilde{\beta}_0$ remain unchanged. If one is employing multiple emission lines to estimate the BHMF, then we have found it faster to simulate proposed values of β_0 for each emission line simultaneously from a multivariate normal distribution.
5. Draw a proposal value for $\log \sigma_{BL}^2$ as $\log \tilde{\sigma}_{BL}^2 \sim N(2 \log \sigma_{BL}^*, \sigma_{\sigma_{BL}}^2)$, where $\sigma_{\sigma_{BL}}^2$ is some set variance. Similar to the update for σ_l^2 , calculate the ratio $r_{BL} = \sigma_{BL}^* p(\theta^*|x_{obs}) / \tilde{\sigma}_{BL} p(\tilde{\theta}|x_{obs})$. If $r_{BL} \geq 1$ then set $\tilde{\sigma}_{BL} = \sigma_{BL}^*$, otherwise set $\tilde{\sigma}_{BL} = \sigma_{BL}^*$ with probability r_σ . This is done by drawing a uniformly distributed random variable as in step 2. If one is employing multiple emission lines to estimate the BHMF, then we have found it faster to simulate proposed values of $\log \sigma_{BL}^2$ for each emission line simultaneously from a multivariate normal distribution.
6. Draw new values of the Gaussian mixture model parameters according to steps 2–6 in the MHA described in KJV08.

One then repeats steps 2–6 until the MCMC converges, saving the values of $\tilde{\theta}$ at each iteration. After convergence, the MCMC is stopped, and the values of $\tilde{\theta}$ may be treated as a random draw from the marginal posterior distribution of θ , $p(\theta|x_{obs})$. Techniques for monitoring convergence of the Markov Chains can be found in Gelman et al. (2004). Given the values of θ obtained from the MCMC, one can then draw values of N from the negative binomial distribution (cf. Eq.[16]).

Having obtained random draws of N and θ from $p(\theta, N|\mathbf{v}_{obs}, L_{obs}, z_{obs})$, one can then use these values to calculate an estimate of $\phi(M_{BH}, z)$, and its corresponding uncertainty. This is done by using each of the MCMC draws of θ and N to calculate Equation (18). The posterior distribution of $\phi(M_{BH}, z)$ can be estimated for any value of M_{BH} and z by plotting a histogram of the values of $\phi(M_{BH}, z)$ obtained from the MCMC values of θ and N . KJV08 illustrates in more detail how to use the MHA results to perform statistical inference.

6. APPLICATION TO SIMULATED DATA

As an illustration of the effectiveness of our method, we applied it to a simulated data set. Because we will eventually apply this method to the BHMF for the SDSS DR3 quasar catalogue (Schneider et al. 2005), we assume the effective survey area and selection function reported for the DR3 quasar sample (Richards et al. 2006).

6.1. Construction of the Simulated Sample

We construct our simulated survey in a manner very similar to that used by KJV08. We first drew a random value of N_Ω quasars from a binomial distribution with probability of success $\Omega/4\pi = 0.0393$ and number of trials $N = 2 \times 10^5$. Here, $\Omega = 1622 \text{ deg}^2$ is the effective sky area for our simulated survey, and we chose the total number of quasars to be $N = 2 \times 10^5$ in order to produce a value of $n \sim 1000$ observed sources after including the flux limit. While this produces a much smaller sample than the actual sample of $\sim 1.5 \times 10^4$ quasars from the SDSS DR3 luminosity function

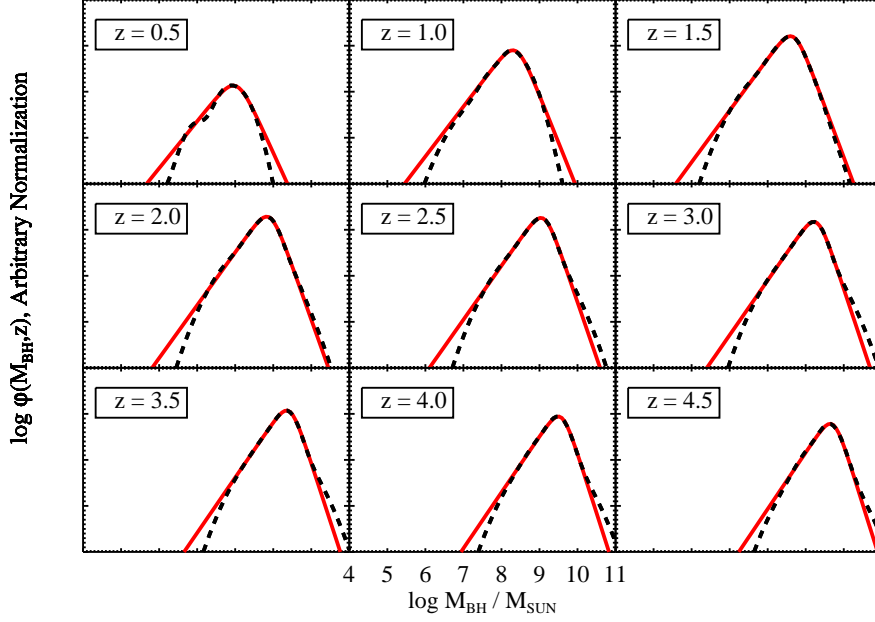


FIG. 2.— The true BHMf (solid red line) at several values of z , and the best $K = 4$ Gaussian function fit (dashed black line). In this case, approximating the BHMf with $K = 4$ 2-dimensional Gaussian functions provides a good fit. The mixture of Gaussian functions approximation diverges from the true BHMf in the tails of the distribution of M_{BH} . However, in general, the uncertainties on the BHMf in the tails are dominated by the statistical errors due to the small number of sources in these regions, and not by the bias introduced from approximating the BHMf as a mixture of Gaussian functions.

work (Richards et al. 2006), we chose to work with this smaller sample to illustrate the effectiveness of our method on more moderate sample sizes. This first step of drawing from a binomial distribution simulates a subset of N_Ω sources randomly falling within an area Ω on the sky, where the total number of sources is N . Note that we have not included any flux limits yet.

For each of these $N_\Omega \sim 8000$ sources, we simulated values of M_{BH} and z . We first simulated values of $\log z$ from a distribution of the form

$$g(\log z) = \frac{4\Gamma(a+b)}{\Gamma(a)\Gamma(b)} \frac{\exp(a\zeta^*)}{(1 + \exp(\zeta^*))^{a+b}}, \quad (60)$$

where $\zeta^* = 4(\log z - 0.4)$. The parameters $a = 2$ and $b = 3$ were chosen to give an observed redshift distribution similar to that seen for SDSS DR3 quasars (e.g., Richards et al. 2006).

For each simulated value of z , we simulated a value of M_{BH} by taking the distribution of M_{BH} at a given redshift to be a smoothly-connected double power-law. In this case, the conditional distribution of $\log M_{BH}$ at a given z is

$$g(\log M_{BH}|z) \propto M_{BH}^{\gamma(z)/\ln 10} \left[1 + \left(\frac{M_{BH}}{M_{BH}^*(z)} \right)^{(\gamma(z)+\delta(z))/\ln 10} \right]^{-1} \quad (61)$$

$$\gamma(z) = 2.5 + 0.5 \log z \quad (62)$$

$$\delta(z) = 4.75 + 2 \log z \quad (63)$$

$$\log M_{BH}^*(z) = 7.5 + 3 \log(1 + z), \quad (64)$$

where $\log M_{BH}^*(z)$ approximately marks the location of the peak in $g(\log M_{BH}|z)$, $\gamma(z)$ is the slope of $\log g(\log M_{BH}|z)$ for $M_{BH} \lesssim M_{BH}^*(z)$, and $\delta(z)$ is the slope of $\log g(\log M_{BH}|z)$ for $M_{BH} \gtrsim M_{BH}^*(z)$. For our simulation, both the peak and logarithmic slopes of the BHMf evolve.

The joint probability distribution of $\log M_{BH}$ and $\log z$ is $g(\log M_{BH}, \log z) = g(\log M_{BH}|z)g(\log z)$, and therefore Equations (60) and (61) imply that the true BHMf for our simulated sample is

$$\phi_0(M_{BH}, z) \propto \frac{N}{z M_{BH}} \left(\frac{dV}{dz} \right)^{-1} g(\log M_{BH}|z)g(\log z). \quad (65)$$

The constant of proportionality in Equation (65) can be calculated by noting that $\int \int \phi_0(M_{BH}, z) dM_{BH} dV = N$. Figure 2 shows $\phi_0(M_{BH}, z)$ at several redshifts. Also shown in Figure 2 is the best fit for a mixture of $K = 4$ Gaussian functions. Despite the fact that $\phi_0(M_{BH}, z)$ has a rather complicated parametric form, a mixture of four Gaussian functions is sufficient to achieve a good approximation to $\phi_0(M_{BH}, z)$.

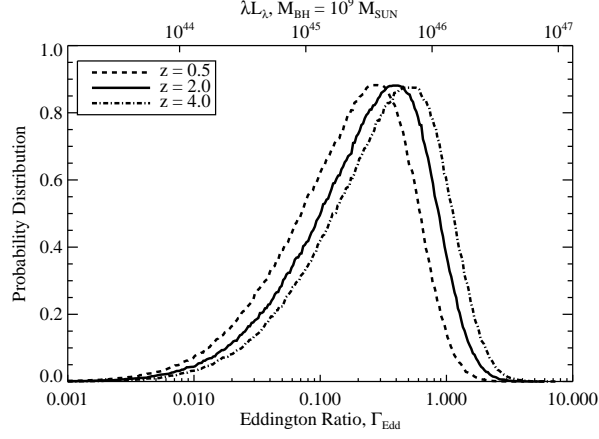


FIG. 3.— Distribution of Eddington ratios, Γ_{Edd} , for our simulated survey at $z = 0.5, 2$, and 4 . The corresponding values of $\lambda L_\lambda[2500\text{\AA}]$ are shown along the top of the plot for a black hole with $M_{BH} = 10^9 M_\odot$ and a bolometric correction of $C_\lambda = 10$.

For each simulated black hole mass and redshift, we simulated a luminosity according to Equation (20). However, unlike the Gaussian distribution assumed in this work (see Eq.[19]), we assume an asymmetric distribution of Eddington ratios that evolves as $\bar{\Gamma}_{Edd} \propto \sqrt{1+z}$. We do this in order to test the robustness of our simple assumption that the distribution of L_λ at a given M_{BH} is independent of redshift and given by a normal distribution. In this simulated ‘universe’, the distribution of Γ_{Edd} does not evolve strongly, as is implied by observations (e.g., Vestergaard 2004; Kollmeier et al. 2006).

To simulate values of luminosity at a given black hole mass, we first simulated values of the Eddington ratio from a skew-normal distribution as

$$\log \Gamma_{Edd} = 0.2\epsilon - 0.75|\delta| - 0.3 + 0.5 \log(1+z). \quad (66)$$

Here, ϵ and δ are both random deviates independently drawn from the standard normal distribution, i.e., $\epsilon, \delta \sim N(0, 1)$. Figure 3 shows the distribution of Γ_{Edd} at a few different redshifts. Values of λL_λ were then calculated according to Equation (20) assuming a constant bolometric correction of $C_\lambda = 10$ (e.g., Kaspi et al. 2000). For simplicity, we only use a constant bolometric correction for all simulated quasars. In our simulation we take $\lambda = 2500\text{\AA}$; the choice of λ is arbitrary and has no material effect on our results. The median Eddington ratio for our simulated sample is $\Gamma_{Edd} \approx 0.25$, and the dispersion in $\log \Gamma_{Edd}$ is ≈ 0.5 dex. Because the mean Γ_{Edd} evolves in our simulation, and because the mean M_{BH} evolves, Γ_{Edd} and M_{BH} are slightly correlated due to the shared correlation with z : $\Gamma_{Edd} \propto M_{BH}^{0.09}$. Therefore, $L_\lambda \propto M_{BH}^{1.09}$. Comparison with Equation (19) suggest that we would expect $\alpha_0 \sim 36$, $\alpha_m \sim 1.09$, and $\sigma_l \sim 0.5$ dex.

For each simulated black hole mass and luminosity, we simulated broad emission line widths for H β , Mg II, and C IV according to Equation (24). We simulated values of the H β line width for $0 < z < 0.9$, values of the Mg II line width for $0.4 < z < 2.2$, and values of the C IV line width for $1.6 < z < 4.5$. Note that for this simulation both H β and Mg II are available at $0.4 < z < 0.9$, and both Mg II and C IV are available at $1.6 < z < 2.2$. Based on the most recent reverberation mapping data (Kaspi et al. 2005; Bentz et al. 2006), we took $R \propto L_\lambda^{1/2}$ ($\beta_l = 0.5$) for all emission lines. In addition, we set $\beta_0 = 10.6, 10.6$, and 10.7 for the H β , Mg II, and C IV emission lines, respectively; these values were chosen to give emission line *FWHM* with typical values of several thousand km s^{-1} . The dispersion in the logarithm of the emission line width at a given luminosity and black hole mass was taken to be $\sigma_{BL} = 0.25, 0.225$, and 0.2 for H β , Mg II, and C IV, respectively. These values of σ_{BL} were chosen to give broad line mass estimate statistical uncertainties similar to that found from the reverberation mapping data (Vestergaard & Peterson 2006).

We randomly kept each source, where the probability of including a source given its luminosity and redshift was taken to be the SDSS DR3 Quasar selection function, as reported by Richards et al. (2006). In addition, we only kept sources with at least one emission line having a line width $1000 \text{ km s}^{-1} < v < 1.8 \times 10^4 \text{ km s}^{-1}$. Sources with $v < 1000$ were assumed to be indistinguishable from narrow-line AGN, and sources with $v > 1.8 \times 10^4$ were assumed to be too difficult to distinguish from the underlying continuum and iron emission, and are thus too broad to be able to obtain a reliable estimate of the line width. After simulating the effects of the selection function, we were left with a sample of $n \sim 1000$ sources. Therefore, our simulated survey was only able to detect $\sim 0.5\%$ of the $N = 2 \times 10^5$ total quasars in our simulated ‘universe’.

The distributions of M_{BH} , z , L_λ , and v are shown in Figure 4 for both the detected sources and the full sample. As can be seen, the majority of sources are missed by our simulated survey, and that the fairly ‘hard’ limit on luminosity corresponds to a much ‘softer’ limit on M_{BH} . In particular, almost all simulated quasars with $M_{BH} \lesssim 10^8 M_\odot$ are missed at $z \gtrsim 1$, and all simulated quasars with $M_{BH} \lesssim 10^7 M_\odot$ are missed at any redshift.

To simulate the effects of using values of β_0 and σ_{BL} derived from a reverberation mapping sample, we simulated a sample of 25 low- z sources with known M_{BH} ; these low- z sources were simulated in the same manner as described above. We then used these 25 ‘reverberation mapping’ sources to fit β_0 and σ_{BL} . The fitted values were then used for our prior distribution on β_0 and σ_{BL} as described in § 4.3.

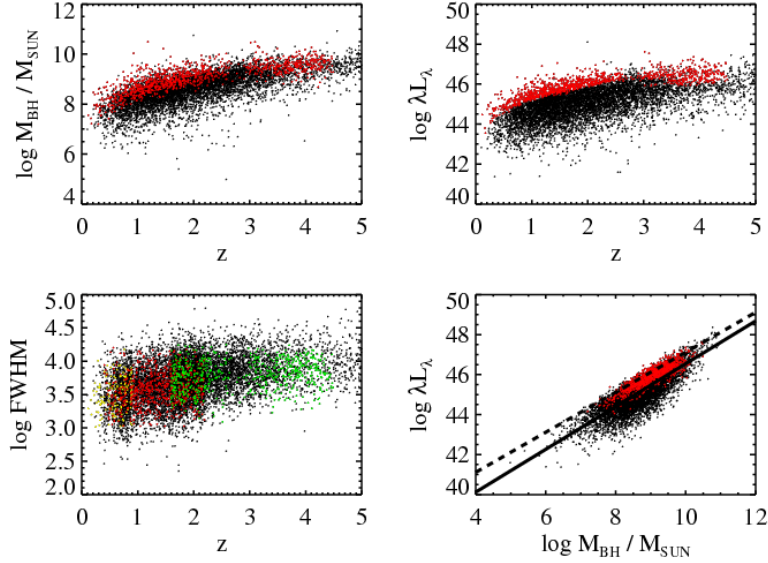


FIG. 4.— The distribution of M_{BH} , L_λ , and $FWHM$ for our simulated sample. Red dots denote sources included in the sample, and black dots denote sources not included in the sample. In the plot of $FWHM$ as a function of z , yellow dots denote sources with $H\beta$ measurements, red dots denote sources with $Mg\ II$ measurements, and green dots denote sources with $C\ IV$ measurements. In the plot of L_λ as a function of M_{BH} , the solid line shows the best linear regression of $\log L_\lambda$ as a function of $\log M_{BH}$, and the dashed line shows the Eddington limit for a bolometric correction of $C_\lambda = 10$.

6.2. Performing Statistical Inference on the BHMF with the MCMC Output

We performed the MHA algorithm described in § 5 to obtain random draws from the posterior distribution for this sample, assuming the Gaussian mixture model described in § 4. We performed 10^4 iterations of burn-in, and then ran the markov chains for an additional 3×10^4 . We ran five chains at the same time in order to monitor convergence (e.g., see Gelman et al. 2004) and explore possible multimodality in the posterior. The chains had converged after 4×10^4 total iterations, leaving us with $\sim 1.5 \times 10^5$ random draws from the posterior distribution, $p(\theta, N | \mathbf{v}_{obs}, L_{obs}, z_{obs})$.

In Figure 5 we show $\phi(\log M_{BH}, z)$ at several different redshifts, on both a linear scale and a logarithmic scale. In general, we find it easier to work with $\phi(\log M_{BH}, z) = \ln 10 M_{BH} \phi(M_{BH}, z)$, as $\phi(M_{BH}, z)$ can span several orders of magnitude in M_{BH} . Figure 5 shows the true value of the BHMF, $\phi_0(\log M_{BH}, z)$, the best-fit estimate of $\phi(\log M_{BH}, z)$ based on the mixture of Gaussian functions model, and the regions containing 68% of the posterior probability. Here, as well as throughout this work, we will consider the posterior median of any quantity to be the ‘best-fit’ for that quantity. In addition, in this work we will report errors at the 68% level unless specified otherwise, and therefore the regions containing 68% of the posterior probability can be loosely interpreted as asymmetric error bars of length $\approx 1\sigma$. As can be seen, the true value of $\phi(\log M_{BH}, z)$ is contained within the 68% probability region for most of the values of $\log M_{BH}$, even those below the survey detection limit.

We compare our method with an estimate of the BHMF obtained by combining the broad line mass estimates with the more traditional $1/V_a$ estimator, developed for luminosity function estimation. We do this primarily to illustrate the pitfalls that can arise from employing broad line mass estimates and not properly accounting for the black hole mass selection function. Following Fan et al. (2001), we denote the effective volume of the i^{th} source as V_a^i . If the i^{th} source lies in a redshift bin of width Δz and has a luminosity $L_{\lambda,i}$, then

$$V_a^i = \int_{\Delta z} s(L_{\lambda,i}, z) \left(\frac{dV}{dz} \right) dz. \quad (67)$$

Dividing up the $(\log M_{BH}, z)$ plane into bins of width $\Delta \log M_{BH} \times \Delta z$, one may be tempted to calculate an estimate of $\phi(\log M_{BH}, z)$ based on the broad line estimates of $\log M_{BH}$ as

$$\hat{\phi}_{BL}(\log M_{BH}, z) = \frac{1}{\Delta \log M_{BH}} \sum_i \frac{1}{V_a^i}. \quad (68)$$

Here, the sum is over all sources with broad lines estimates $\log M_{BH} \leq \log \hat{M}_{BL,i} \leq \log M_{BH} + \Delta \log M_{BH}$ and $z \leq z_i \leq z + \Delta z$.

Figure 5 also displays the expected value of $\hat{\phi}_{BL}$ for $z = 0.5, 1.5, 2.5, 3.5$ and 4.5 . In order to estimate the expected value of $\hat{\phi}_{BL}$ at each z , we simulated 10^7 quasars at each redshift interval. This produces extremely small error bars on $\hat{\phi}_{BL}$ and allows us to estimate the value of $\hat{\phi}_{BL}$ that would be obtained on average, i.e., in the limit of an infinitely large sample. As can be seen, $\hat{\phi}_{BL}$ is a biased estimate of the BHMF. This bias is caused by a combination of the relatively large statistical uncertainties on the broad line mass estimates, which broaden the inferred BHMF, and by the use of the luminosity selection function instead of the black hole mass selection function in the $1/V_a$ correction.

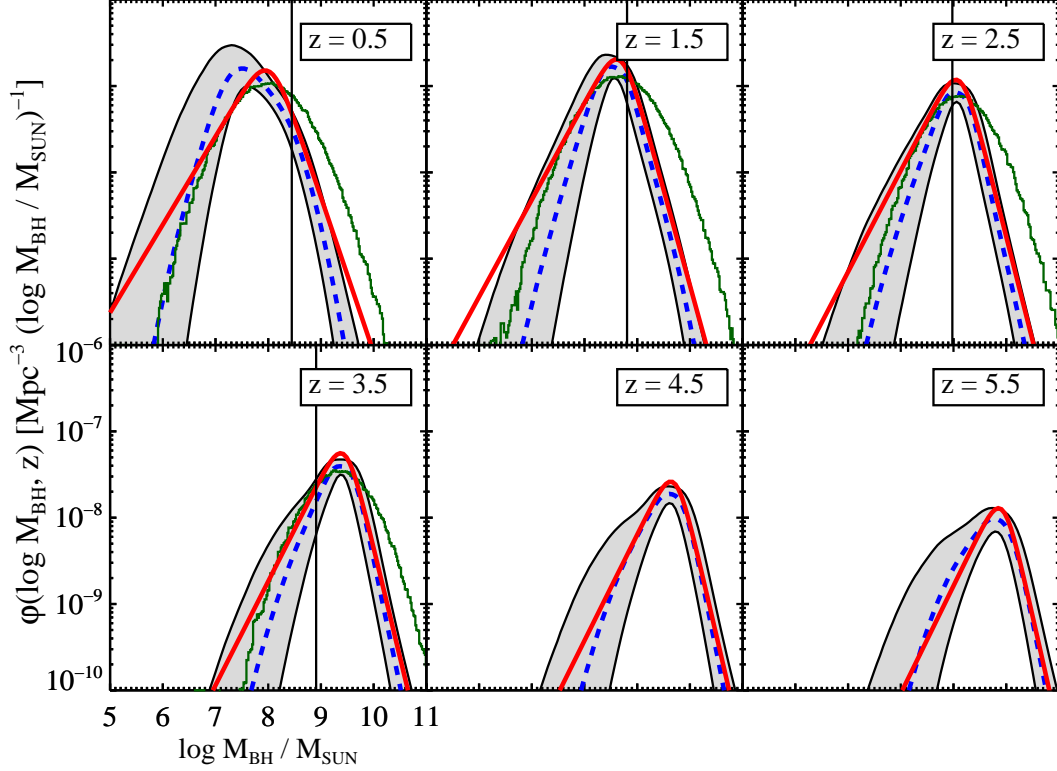


FIG. 5.— The true BHMf (solid red line) at several redshifts. The axis labels are the same for all panels, but for clarity we only place exterior labels on the bottom left panel. Also shown is the posterior median estimate of the BHMf based on the mixture of Gaussian functions model (dashed blue line), the region containing 68% of the posterior probability (shaded region), and the expected value for a $1/V_a$ -type binned estimate based on the broad emission line estimates, $\hat{\phi}_{BL}$ (thin bumpy solid green line). The vertical lines mark the 50% incompleteness limit for a quasar with $\text{FWHM} = 4000 \text{ km s}^{-1}$, a typical value for the simulated sources. Note that in general the best-fit mixture of Gaussian functions approximation to the BHMf will not equal the true BHMf, as it is derived from a finite random sample drawn from the true BHMf. The Bayesian mixture of Gaussian functions model is able to accurately constrain the BHMf, even below the survey detection limit. However, $\hat{\phi}_{BL}$ provides a biased estimate of the BHMf.

The large statistical uncertainties on the broad line mass estimates broaden the inferred BHMf, and therefore $\hat{\phi}_{BL}$ significantly overestimates the BHMf at the high mass end, while underestimating the BHMf near its peak. In addition, $\hat{\phi}_{BL}$ underestimates the BHMf at the low mass end due to the inability of the $1/V_a$ technique to completely correct for incompleteness. The end result is a systematic shift in the inferred BHMf toward higher M_{BH} , and a similar effect has been noted by Shen et al. (2007). The effective volume in Equation (67) is defined based on the detection probability as a function of luminosity, and not black hole mass. As mentioned in § 2.3, in order to correctly apply the $1/V_a$ estimator for BHMf estimation it is necessary to obtain the black hole mass selection function, given by Equation (11). However, this requires knowledge of $p(L_\lambda | M_{BH}, z)$. Furthermore, even if there were no selection effects, $\hat{\phi}_{BL}$ would still be biased because of the significant uncertainty ($\sim 0.4 \text{ dex}$) on $\log \hat{M}_{BL}$.

As in KfV08, we can use the MCMC output to constrain various quantities of interest calculated from the BHMf. Figure 6 compares the true integrated $z < 6$ number distribution of $\log M_{BH}$, $n(\log M_{BH}, z < 6)$, with the mixture of Gaussian functions estimate. The quantity $n(\log M_{BH}, z < 6) d \log M_{BH}$ is the number of quasars at $z < 6$ with black hole masses between $\log M_{BH}$ and $\log M_{BH} + d \log M_{BH}$. KfV08 give an equation for calculating $n(\log L, z < z_0)$ based on the mixture of Gaussian functions model (see their Eq.[42]), and $n(\log M_{BH}, z < z_0)$ is calculated in an equivalent manner. Similar to Figure 5, the true value of $n(\log M_{BH}, z < 6)$ is contained within the 68% probability region for most values of M_{BH} , even those below the survey detection limit.

In addition, in Figure 6 we show the comoving number density of broad line AGN as a function of redshift, $n(z)$. This is obtained by integrating $\phi(M_{BH}, z)$ over all possible values of M_{BH} , given by Equation (45) of KfV08. As before, the true value of $n(z)$ is contained within the 68% probability region, despite the fact that the integration extends over *all* M_{BH} , even those below the detection limit. The wider confidence regions reflect additional uncertainty in $n(z)$ resulting from integration over those M_{BH} below the detection limit. In particular, the term dV/dz becomes small at low redshift, making the estimate of $n(z)$ more unstable as $z \rightarrow 0$, and thus inflating the uncertainties at low z .

Two other potentially useful quantities are the comoving black hole mass density for quasars, $\rho_{BH}^{QSO}(z)$, and its derivative. The comoving black hole mass density is given by $\rho_{BH}^{QSO}(z) = \int_0^\infty M_{BH} \phi(M_{BH}, z) dM_{BH}$. The quantity $\rho_{BH}^{QSO}(z)$ is given by Equation (47) of KfV08 and replacing luminosity with black hole mass. We calculate the derivative

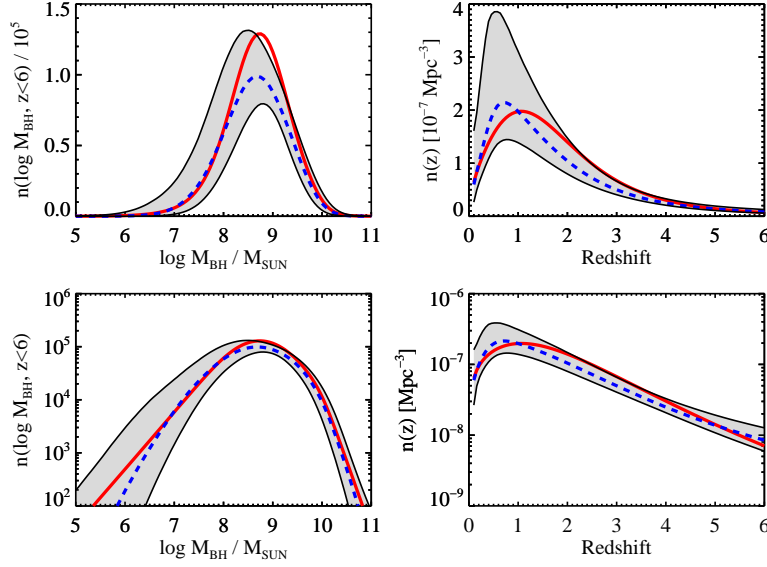


FIG. 6.— The integrated $z < 6$ quasar number density (number per $\log M_{BH}$ interval, left two panels) and the comoving quasar number density as a function of z (number per Mpc^3 , right two panels). The top two panels show a linear stretch and the bottom two panels show a logarithmic stretch. As with Figure 5, the solid red line denotes the true value for the simulation, the dashed blue line denotes the posterior median for the mixture of Gaussian functions model, and the shaded regions contain 68% of the posterior probability. The posterior median provides a good fit to the true values, and the uncertainties derived from the MCMC algorithm based on the Gaussian mixture model are able to accurately constrain the true values of these quantities, despite the flux limit.

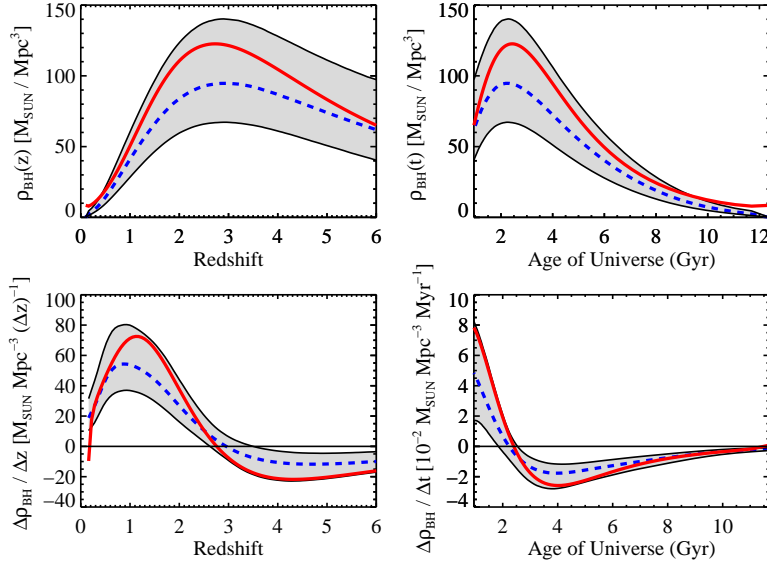


FIG. 7.— Comoving broad line quasar black hole mass density (top two panels) and its derivative (bottom two panels), shown as a function of redshift (left two panels) and cosmic age (right two panels). The plotting symbols are the same as in Figure 6. As in the previous figures, the Gaussian mixture model is able to provide an accurate fit to the true values of $\rho_{BH}^{QSO}(z)$, and the bayesian MCMC approach is able to provide accurate constraints on $\rho_{BH}^{QSO}(z)$ and $d\rho_{BH}^{QSO}/dz$, despite the fact that the integral used for calculating these quantities extends below the survey detection limit.

of $\rho_{BH}^{QSO}(z)$ numerically. Figure 7 compares the true values of $\rho_{BH}^{QSO}(z)$ and its derivative with the posterior distribution for $\rho_{BH}^{QSO}(z)$ inferred from the mixture model, both as a function of z and the age of the universe at redshift z , $t(z)$. Comparison with Figure 6 reveals that the comoving quasar black hole mass density, $\rho_{BH}^{QSO}(z)$, is a better constrained quantity than the comoving quasar number density, $n(z)$. Furthermore, $n(z)$ appears to peak later than $\rho_{BH}^{QSO}(z)$. We can correctly infer that the quasar comoving black hole mass density reaches its point of fastest growth at $t(z) \lesssim 1$ Gyr, and its point of fastest decline at $t(z) \sim 4$ Gyr.

Figure 8 quantifies the suggestion that $n(z)$ peaks later than $\rho_{BH}^{QSO}(z)$ by displaying the posterior distribution for the location of the respective peaks in $n(z)$ and $\rho_{BH}^{QSO}(z)$. While the location of the peak in $n(z)$ is highly uncertain

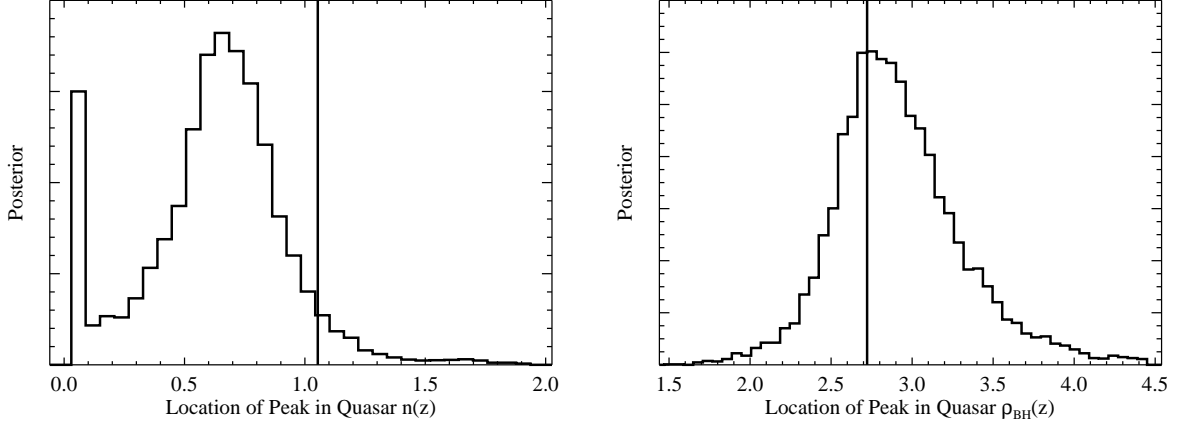


FIG. 8.— Posterior distribution for the redshift location of the peak in the comoving number density of quasars ($n(z)$, left) and the peak in the comoving quasar black hole mass density ($\rho_{BH}^{QSO}(z)$, right). The spike in the posterior at $z \approx 0$ for values of the peak in $n(z)$ arises because the term $(dV/dz)^{-1}$ becomes very large at low z . The vertical lines denote the true values. The posterior distribution inferred from the MCMC output is able to accurately constrain the true values of the argumentative maximum in $n(z)$ and $\rho_{BH}^{QSO}(z)$.

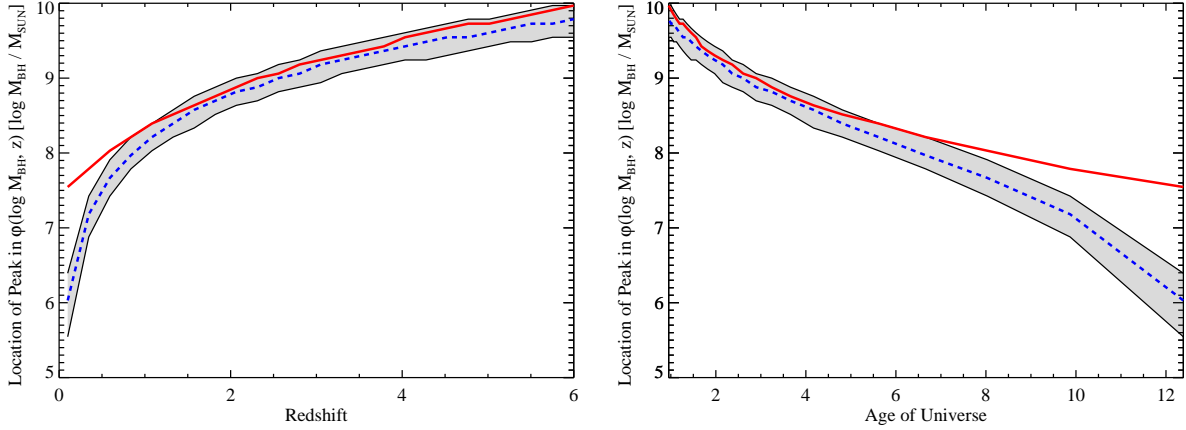


FIG. 9.— Location of the peak in the BHMf as a function of z (left) and cosmic age (right). The plot symbols are the same as in Figure 6. In general the posterior median of the Gaussian mixture model provides a good estimate of the true peak locations, although some bias is exhibited at the lowest redshifts. It is clear from these plots that the location of the peak in $\phi(M_{BH}, z)$ evolves.

we can still constrain it to be $z \lesssim 1.5$, whereas the location of the peak in $\rho_{BH}^{QSO}(z)$ is constrained to occur earlier at $2 \lesssim z \lesssim 4$. This is a consequence of the fact that while there were more quasars at $z \sim 1$ per comoving volume, their black hole masses were much higher at higher redshift. This evolution in characteristic M_{BH} is quantified in Figure 9, which summarizes the posterior distribution for the location of the peak in $\phi(\log M_{BH}, z)$ as a function of redshift and $t(z)$. As can be seen, the location of the peak in the BHMf shows a clear trend of increasing ‘characteristic’ M_{BH} with increasing z , although the mixture of Gaussian functions fit has difficulty constraining the location of the peak at low redshift.

As noted in § 4.2, we can use the values of α_0 and σ_l to estimate the average Eddington ratio and the dispersion in $\log \Gamma_{Edd}$. We find $\alpha_0 = 35.7^{+0.9}_{-1.1}$, $\alpha_m = 1.11^{+0.12}_{-0.10}$, and $\sigma_l = 0.31^{+0.06}_{-0.05}$, where the errors are at 95% confidence. For a bolometric correction of $C_\lambda = 10$, and assuming that Γ_{Edd} is independent of M_{BH} , this implies that our inferred typical Eddington ratio is $\Gamma_{Edd} = 0.040^{+0.278}_{-0.036}$ at 95% confidence; the estimated dispersion in $\log \Gamma_{Edd}$ is simply given by σ_l , ~ 0.3 dex. While the typical Eddington ratio that we infer from α_0 is roughly consistent with the actual median Γ_{Edd} of 0.25, our estimated dispersion in Γ_{Edd} underestimates the true value of 0.5 dex. This is because we incorrectly assume that the M_{BH} – L relationship is described by Equation (19). Our inference regarding the Eddington ratio distribution is therefore biased because we assume that the distribution of Γ_{Edd} does not evolve, and that the distribution is Gaussian. In particular, the bias resulting from the assumption of Gaussian dispersion appears to significantly affect the estimated dispersion in $\log \Gamma_{Edd}$ more than the estimated typical value of Γ_{Edd} , at least for our simulation. This is largely because the distribution in Γ_{Edd} is skewed toward lower values of Γ_{Edd} . However, because of the flux limit, sources with low values of Γ_{Edd} are undetectable. Because the dispersion in $\log \Gamma_{Edd}$ is estimated from the detected sources, in combination with the assumption of a Gaussian distribution, Equation (19) is not able to pick up the additional skew at low $\log \Gamma_{Edd}$. As a result, the estimated dispersion in $\log \Gamma_{Edd}$ is underestimated when assuming a Gaussian distribution. We note that this bias is not a feature of our algorithm, but affects any analysis

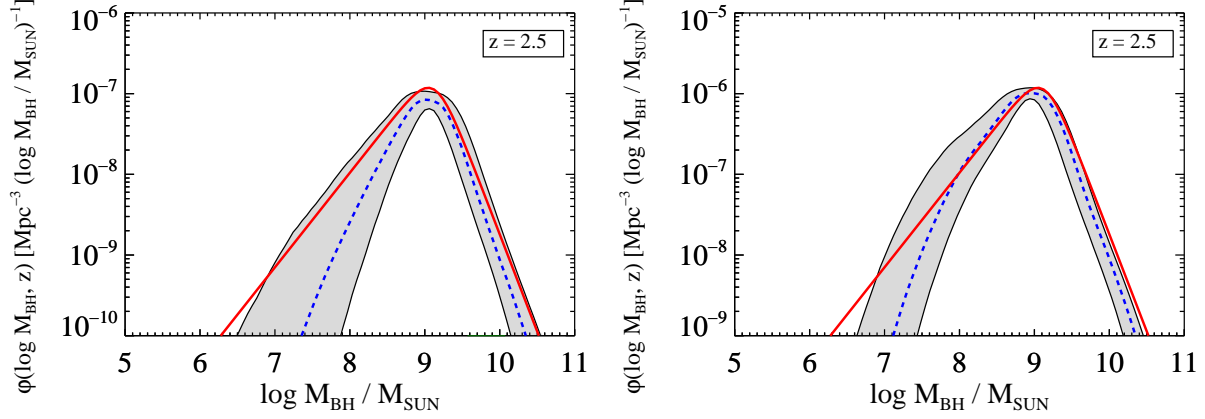


FIG. 10.— BHMf at $z = 2.5$ for the simulated sample with $n \sim 1000$ detected sources (left) and $n \sim 10^4$ detected sources (right); the left panel is the same as the $z = 2.5$ BHMf shown in Figure 5. The uncertainties derived for the $n \sim 10^4$ are smaller than for the $n \sim 1000$ sample, particularly at low M_{BH} where the survey becomes incomplete. However, the uncertainties for the $n \sim 10^4$ survey at high M_{BH} , where the survey is complete, are not considerably smaller than those for the $n \sim 1000$ survey. This is because the BHMf estimate is limited by the systematic uncertainty in the broad line mass estimate normalization, derived from β_0 , and the broad line mass estimate statistical error, derived from σ_{BL} . Because the observed distribution of luminosities and line widths does not convey any information on these two quantities, increasing the sample size will not reduce the uncertainty on the BHMf beyond the systematic uncertainty on β_0 and σ_{BL} .

that attempts to infer the distribution of Eddington ratios using a flux-limited sample.

In order to assess how the inferred BHMf depends on the sample size, we simulated a second data set in the same manner as described above, but used $N = 2 \times 10^6$ sources for the BHMf normalization. This gave us $n \sim 10^4$ detected quasars. In Figure 10 we compare the estimated BHMf at $z = 2.5$ for the survey with $n \sim 1000$ sources and $n \sim 10^4$ sources. The uncertainties are lower for the survey with more sources, where the most noticeable improvement occurs at low M_{BH} . However, the increased sample size did not offer a significant amount of improvement at high M_{BH} , where sources are more easily detected. This is likely because the uncertainty in the broad line mass estimate normalization, β_0 , and intrinsic scatter, σ_{BL} , dominates the uncertainty in the BHMf at high M_{BH} . Because we cannot constrain β_0 and σ_{BL} from the distribution of line widths and luminosities, the data do not contain any information on β_0 and σ_{BL} . Therefore, the likelihood function is unable to convey any information on β_0 and σ_{BL} , and all of the information comes from the prior distribution. As a result, our ability to constrain the BHMf is limited by the statistical uncertainty on β_0 and σ_{BL} , and an increase in the sample size will eventually not result in a decrease in the uncertainty on the BHMf. The only way to reduce the uncertainty on the BHMf for large surveys is to better constrain the broad line mass estimate normalization and statistical uncertainty, most likely by increasing the sample of AGN with reverberation mapping data.

Throughout this work we have assumed that the selection function is known, and that β_0 , and σ_{BL} are known within some statistical uncertainty. However, this may not be the case, and before concluding this section we briefly discuss how systematic error in the selection function, β_0 and σ_{BL} , affect the inferred BHMf. We did not experiment with incorrect selection functions, and so it is not entirely clear how robust BHMf estimation is to errors in the selection function. However, from Equations (15) and (16) it is clear that the selection function only enters into the posterior probability distribution (or likelihood function) via an integral that averages the selection function over the joint distribution of luminosity and redshift (i.e., the luminosity function). As a result, errors in the selection function will be smoothed out. Furthermore, they will be suppressed in regions where values of the luminosity function are small. Based on this, we do not think it likely that small errors in the selection function will introduce significant bias into the results; however, if the errors in the selection function are large enough to significantly bias the value of $p(I = 1|L, z)$, then the results may be significantly biased as well.

It is useful to work directly with the broad line mass estimates to assess the effect that systematic uncertainty on the values of the broad line mass estimate normalization and statistical uncertainty have on the inferred BHMf. Ignoring selection effects, one can think of our method as ‘correcting’ the BHMf inferred from binning up the broad line mass estimates. Therefore, if β_0 is systematically underestimated, then this will result in a shift of the inferred BHMf toward higher masses. Similarly, if β_0 is systematically overestimated, then the inferred BHMf will be shifted toward lower masses. In addition, the value of σ_{BL} controls how much the BHMf inferred from the broad line mass estimates is artificially broadened by the statistical uncertainty in \hat{M}_{BL} . A higher value of σ_{BL} will result in a greater amount of broadening. Therefore, if our assumed values of σ_{BL} are systematically overestimated, then we would infer a greater amount of broadening than is real. As a result, our correction would be too large, and we would infer an intrinsic BHMf that is too narrow. Similarly, if our assumed values of σ_{BL} are systematically underestimated, then we would not correct enough for the statistical uncertainty in the broad line mass estimates, and we would infer an intrinsic BHMf that is too broad.

6.3. Using the MCMC Output to Evaluate the BHMf Fit

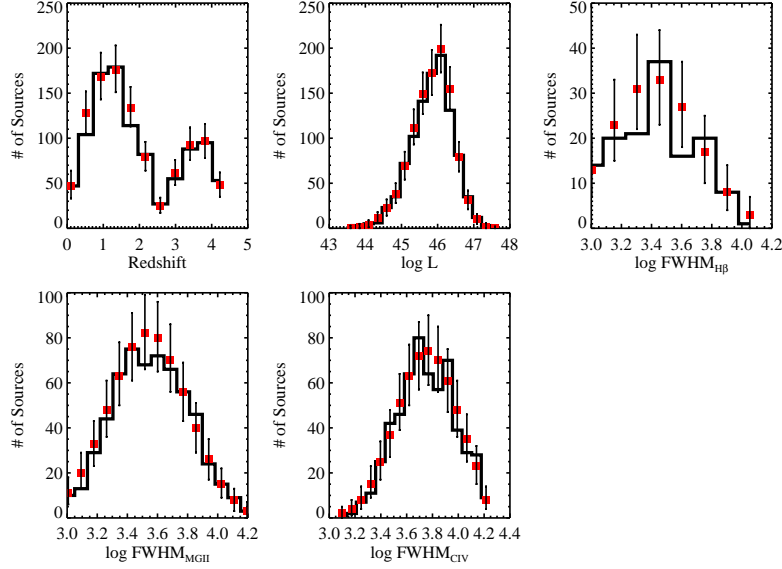


FIG. 11.— Posterior predictive check for the Gaussian mixture model (see § 6.3). The histograms show the actual distributions of $\log L_{obs}$, z_{obs} , and $\log FWHM_{obs}$, the red squares denote the posterior medians for the number of sources in each respective bin, and the error bars contain the inner 90% of the histogram values for the samples simulated from the posterior. The mixture of Gaussian functions model is able to provide an accurate prediction of the observed distribution of luminosity, redshift, and line widths, and thus there is not any evidence to reject it as providing a poor fit.

Throughout this section we have been analyzing the MCMC results by comparing to the true BHMF. However, in practice we do not have access to the true BHMF, and thus a method is needed for assessing the quality of the fit. As in KJV08, the statistical model may be checked using a technique known as posterior predictive checking (e.g., Rubin 1981, 1984; Gelman, Meng, & Stern 1998). Here, the basic idea is to use each of the MCMC outputs to simulate a new random observed data set. The distributions of the simulated observed data sets are then compared to the true observed data in order to assess whether the statistical model gives an accurate representation of the observed data. It is important to construct simulated data sets for each of the MCMC draws in order to incorporate our uncertainty in the model parameters.

Random draws for M_{BH} and z for each MCMC draw may be obtained according to the procedure outlined in § 7.3 of KJV08, after replacing L with M_{BH} . Once one obtains a random draw of M_{BH} and z , simulated values of L_λ may be obtained using Equation (19) with α_0 , α_m , and σ_l . Then, given these values of L_λ and M_{BH} , values of v for each emission line can be simulated from Equation (24) using the values of β_0 , β_l , and σ_{BL} . Simulation from Equation (24) requires a value of α_λ in order to convert L_λ to L_λ^{BL} . In order to account for the range in continuum slopes, we randomly draw a value of α_λ from our data set and use this value to convert to L_λ^{BL} . These simulated values of L_λ , z , and \mathbf{v} are then folded through the selection function, leaving one with a simulated observed data set $(\mathbf{v}_{obs}, L_{obs}, z_{obs})$. This process is repeated for all values of N and θ obtained from the MCMC output, leaving one with simulated observed data sets of $(\mathbf{v}_{obs}, L_{obs}, z_{obs})$. These simulated observed data sets can then be compared with the true distribution of \mathbf{v}_{obs} , L_{obs} , and z_{obs} to test the statistical model for any inconsistencies.

In Figure 11 we show histograms for the observed distributions of z , $\log L_\lambda$, and $\log FWHM$ for the H β , Mg II, and C IV emission lines. These histograms are compared with the posterior median of the observed distributions based on the mixture of Gaussian functions model, as well as error bars containing 90% of the simulated observed values. As can be seen, the distributions of the observed data sets simulated from our assumed statistical model are consistent with the distributions of the true observed data, and therefore there is no reason to reject the statistical model as providing a poor fit.

7. APPLICATION TO BQS QUASARS

As a final illustration of our method we used it to estimate the low redshift active BHMF from the 87 $z < 0.5$ quasars from the Bright Quasar Survey (BQS, Schmidt & Green 1983). The H β line widths and continuum luminosities for 71 of the BQS quasars are taken from Table 7 of Vestergaard & Peterson (2006), and 16 of the quasars in the Boroson & Green (1992) sample have black hole mass estimates from reverberation mapping (Peterson et al. 2004). For each source with reverberation mapping data, we used the first entry of $\lambda L_\lambda(5100\text{\AA})$ in Table 1 of Vestergaard & Peterson (2006) as the single-epoch luminosity; these values were based on continuum luminosities reported by Boroson & Green (1992) or Marziani et al. (2003). We assumed measurement errors of 10% on the emission line $FWHM$. The BQS sample covers an area of $\Omega = 10,714 \text{ deg}^2$ and is selected with an average flux limit of $B = 16.16$ (Schmidt & Green 1983), with no apparent correlation with redshift and $U - B$ color (Jester et al. 2005). We converted the $B = 16.16$ flux limit to a flux limit at 5100\AA assuming a power law continuum, $f_\nu \propto \nu^{-\alpha}$, with $\alpha = 0.5$ (Richards et al. 2001). We used $K = 3$ Gaussian functions to fit $\phi(M_{BH}, z)$ for $z < 0.5$.

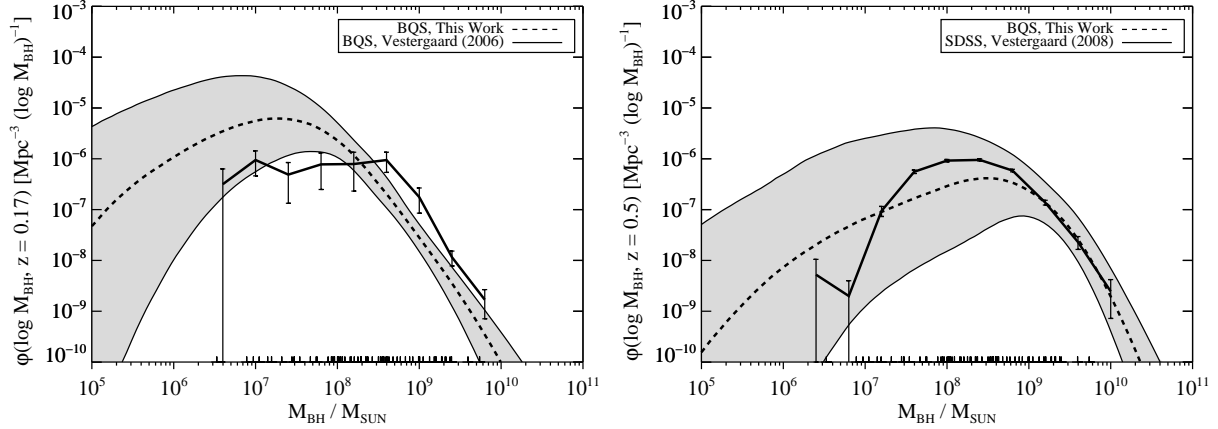


FIG. 12.— The $z = 0.17$ (left) and $z = 0.5$ (right) broad line quasar BQSBHMF as estimated from the BQS sample. The dashed line denotes the posterior median for the mixture of Gaussian functions model, the shaded region contains 68% of the posterior probability, and the tick marks along the x -axis mark the locations of the broad line mass estimates. The estimate of the $z = 0.17$ BQSBHMF becomes significantly uncertain at $M_{BH} \lesssim 10^8 M_\odot$, and the $z = 0.17$ BQSBHMF appears to fall off as a power law above $M_{BH} \gtrsim 10^8 M_\odot$. The $z = 0.5$ BQSBHMF is not very well constrained, but there is evidence for a shift in the BQSBHMF toward higher M_{BH} from $z = 0.17$ to $z = 0.5$. For comparison, we show the BQSBHMF estimated by Vestergaard (2006) using the BQS sources (left, solid line with error bars), and the BQSBHMF estimated by Vestergaard et al. (2008) using the SDSS DR3 quasars (right, solid line with error bars). The shift in the BQSBHMF inferred from the binned mass estimates is apparent in the BQS sample, while the SDSS and BQS $z = 0.5$ BQSBHMF estimates agree fairly well.

Because we are including the actual values of M_{BH} for the 16 reverberation mapping sources, the contribution to the posterior for these sources is

$$p(\theta|M_{BH}, L_\lambda, z) = \prod_{i=1}^{16} p(\log L_{\lambda,i}|M_{BH,i}, \theta) p(\log M_{BH,i}, \log z_i|\theta). \quad (69)$$

Here, $p(L_{\lambda,i}|M_{BH,i}, \theta)$ is given by Equation (19) and $p(\log M_{BH,i}, \log z_i|\theta)$ is given by Equation (17). The product in Equation (69) is only over the quasars with M_{BH} estimated from reverberation mapping, whereas the contribution to the posterior for the BQS sources without reverberation mapping is given by Equation (15). The posterior for the complete BQS sample is then the product of Equation (69) and Equation (15).

In Figure 12 we show the $z = 0.17$ BQSBHMF derived from the BQS sample. Also shown is the binned BQSBHMF for the BQS sources, calculated directly from the broad line mass estimates by Vestergaard (2006). We show the BQSBHMF at $z = 0.17$ because the average redshift of the BQS sources is $z \approx 0.17$, therefore allowing a more direct comparison between the binned BQSBHMF and the BQSBHMF derived using our mixture of Gaussian functions approach. In addition, the uncertainties on our estimated BQSBHMF are smallest at $z \approx 0.17$. We are able to place some constraints on the local BQSBHMF, despite the fact that the BQS sample only contains 87 sources and has a very shallow flux limit. The $z \sim 0.2$ quasar BQSBHMF appears to fall off as a power law above $M_{BH} \gtrsim 10^8 M_\odot$. Unfortunately, our estimate of the local BQSBHMF becomes considerably uncertain below $M_{BH} \lesssim 10^8 M_\odot$, so it is unclear to what degree the power law trend continues below this point. In addition, the binned estimate overestimates the BQSBHMF at the high M_{BH} end due to the intrinsic uncertainty in the broad line mass estimates, and underestimates the BQSBHMF at the low M_{BH} end due to incompleteness, in agreement with our simulations (see § 6.2).

In Figure 12 we also compare our estimate of the BQSBHMF at $z = 0.5$ with the $z = 0.5$ BQSBHMF as reported by Vestergaard et al. (2008). Vestergaard et al. (2008) estimated the $z = 0.5$ BQSBHMF by binning estimates of M_{BH} derived from the H β and Mg II broad emission lines over the redshift range $0.3 < z < 0.68$, using the SDSS DR3 quasar catalogue (Schneider et al. 2005). Despite the differences in approach and survey selection, the two estimates of the $z = 0.5$ BQSBHMF agree fairly well. However, because $z = 0.5$ defines the upper redshift limit of our BQS sample, the uncertainties on the BQSBHMF derived from the BQS quasars are very large. In addition, incompleteness in M_{BH} likely affects the low M_{BH} bins of the Vestergaard et al. (2008), causing the Vestergaard et al. (2008) $z = 0.5$ BQSBHMF to underestimate the true $z = 0.5$ BQSBHMF in these bins, a fact reflected by the larger error bars. However, a direct comparison between our Bayesian approach and the Vestergaard et al. (2008) estimate is difficult, due to the different redshift ranges used to estimate the BQSBHMF, and the different selection methods of the BQS and the SDSS.

Although the BQS has a small sample size and probes a narrow range in z , we can attempt to quantify any evolution in the local BQSBHMF by comparing the ratio of the comoving number density of quasars at two different values of M_{BH} . Comparison of the estimated BQSBHMF at $z = 0.17$ and $z = 0.5$ suggests a shift in the BQSBHMF toward large M_{BH} . In Figure 13 we show the best fit values of the ratio of $\phi(\log M_{BH}, z)$ at $M_{BH} = 5 \times 10^8 M_\odot$ to $\phi(\log M_{BH}, z)$ at $M_{BH} = 5 \times 10^9 M_\odot$ as a function of z , as well as the 68% confidence interval. The logarithm of this ratio gives the slope of a power-law between $M_{BH} = 5 \times 10^8 M_\odot$ and $M_{BH} = 5 \times 10^9 M_\odot$, and therefore allows us to probe evolution in the shape of the quasar BQSBHMF at the high M_{BH} end. In general, the ratio is fairly flat, implying no evolution in the high M_{BH} slope of the BQSBHMF. However, at $z \gtrsim 0.3$ there is marginal evidence for a flattening of the high M_{BH} slope of the BQSBHMF. The values of this ratio imply that the BQSBHMF at the high M_{BH} end falls off as a power-law with slope ~ 2 , although slopes of ~ 1 and ~ 3 are also consistent with the BQS quasars.

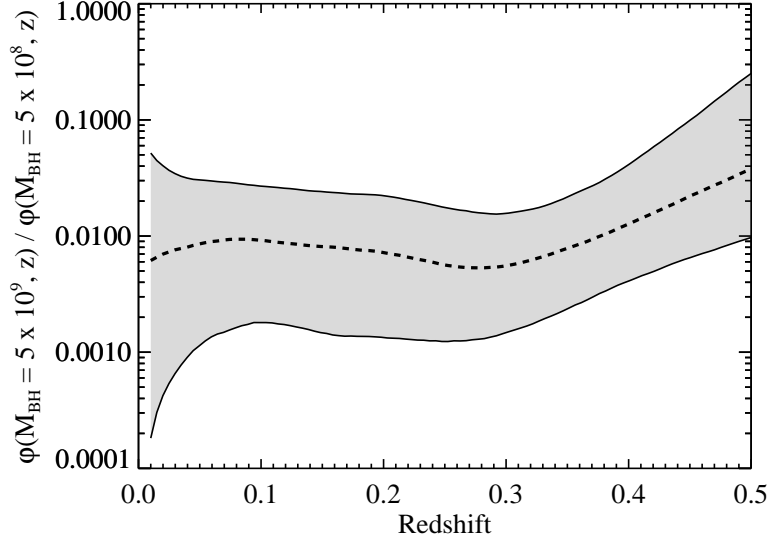


FIG. 13.— The ratio of the broad line quasar BHMf at $M_{BH} = 5 \times 10^9 M_\odot$ compared to the BHMf at $M_{BH} = 5 \times 10^8 M_\odot$, as a function of z and estimated from the BQS quasars. The dashed line is the posterior median, and the shaded region contains 68% of the probability. Assuming that the BHMf is a power-law from $M_{BH} = 5 \times 10^8 M_\odot$ to $M_{BH} = 5 \times 10^9 M_\odot$, the logarithm of this ratio is the slope of the BHMf. The high M_{BH} BHMf slope appears to be fairly constant for $z \lesssim 0.3$ with a slope of ~ 2 , and there is marginal evidence for a flattening of the high M_{BH} slope at $z \gtrsim 0.3$.

In figure 14 we summarize the posterior probability distribution for the parameters governing the distribution of L_λ at a given M_{BH} (see Eq.[19]). Based on the MCMC results, we can constrain the M_{BH} – $\lambda L_\lambda(5100\text{\AA})$ relationship at $z < 0.5$ to be

$$\lambda L_\lambda(5100\text{\AA}) = 5.18^{+429}_{-5.14} \times 10^{36} \left(\frac{M_{BH}}{M_\odot} \right)^{0.92 \pm 0.24} [\text{erg s}^{-1}], \quad (70)$$

where we have quoted the errors at 95% confidence. The dispersion in L_{5100} at a given M_{BH} is estimated to be $\sigma_l = 0.35^{+0.13}_{-0.08}$. Assuming that the bolometric correction is on average $C_{5100} \sim 10$ (e.g., Kaspi et al. 2000), comparison of Equation (70) with Equation (20) suggests that $z < 0.5$ broad line AGN have typical Eddington ratios of $\Gamma_{Edd} \sim 0.4$. As argued in § 4.2, the distribution in $\log L_\lambda$ at a given M_{BH} is the convolution of the distribution of $\log \Gamma_{Edd}$ with the distribution of $\log C_\lambda$. Therefore, the dispersion in L_λ at a given M_{BH} is a combination of the dispersion in Eddington ratio and bolometric correction. As a result, we are unable to estimate the dispersion in Eddington ratios at a given M_{BH} from σ_l . However, if the bolometric correction to L_{5100} increases with increasing Eddington ratio, as found by Vasudevan & Fabian (2007), or if the bolometric correction is independent of Γ_{Edd} , then the dispersion in Γ_{Edd} must be less than σ_l . Therefore, because we infer that $\sigma_l \lesssim 0.5$ dex, our results imply that the dispersion in Eddington ratios at a given M_{BH} is $\lesssim 0.5$ dex for $z < 0.5$ broad line quasars. These results on the Eddington ratio distribution are consistent with previous work (e.g., McLure & Dunlop 2004; Vestergaard 2004; Kollmeier et al. 2006); however, they may be biased because of our assumption of a Gaussian and non-evolving Eddington ratio distribution. In particular, if the distribution of Eddington ratios is skewed toward low $\log \Gamma_{Edd}$, then we will have underestimated the intrinsic dispersion in $\log \Gamma_{Edd}$.

8. SUMMARY

We have derived the observed data likelihood function which relates the quasar BHMf to the observed distribution of redshifts, luminosities, and broad emission line widths. This likelihood function is then used in a Bayesian approach to estimating the BHMf, where the BHMf is approximated as a mixture of Gaussian functions. Because much of this work was mathematically technical, we summarize the important points here.

- In this work we describe a flexible parametric model for the BHMf, where the BHMf is modeled as a mixture of Gaussian functions. The distribution of luminosities is modelled as a linear regression of $\log L_\lambda$ as a function of $\log M_{BH}$, where the distribution of $\log L_\lambda$ at a given M_{BH} was assumed to follow a normal distribution. The distribution in line widths at a given L_λ and M_{BH} is also assumed to have the form of a linear regression, where the parameters are based on the most recent broad line mass estimates. Equation (18) gives the BHMf under the mixture of Gaussian function model.

Equations (30) and (46) define the likelihood function for broad line mass estimates under the mixture of Gaussian functions model if only one emission line at a given z is used to estimate M_{BH} . Otherwise, if multiple emission lines are used for a single quasar, then Equation (41) must be used. The posterior is then found by inserting the prior distribution and likelihood function into Equations (15) and (16).

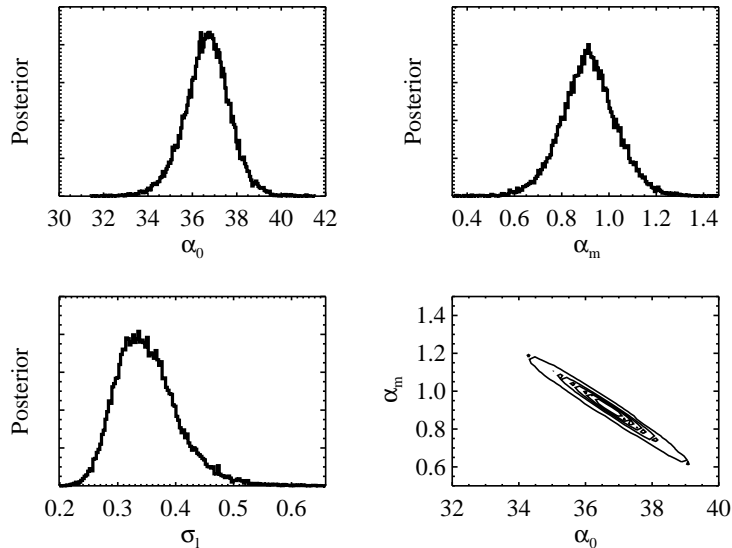


FIG. 14.— Posterior distributions of the parameters for the distribution of luminosities at a given M_{BH} , as estimated from the $z < 0.5$ BQS quasars. The uncertainty on α_0 and α_m is highly correlated. Assuming a bolometric correction of $C_{5100} \sim 10$, the values of α_0 and σ_l imply that the $z < 0.5$ distribution of broad line quasar Eddington ratios has a mean of $\Gamma_{Edd} \sim 0.4$ and a dispersion of ~ 0.5 dex.

- Using methods developed for luminosity function estimation (e.g., $1/V_a$ -type estimators) without modification will lead to errors in black hole mass function estimation, as the black hole mass selection function is not equivalent to the flux selection function. In addition, using broad line estimates of M_{BH} will lead to a broader inferred BHMF if one does not correct for the intrinsic uncertainty in the broad line mass estimates. This causes one to overestimate $\phi(M_{BH}, z)$ in the tails of the distribution, and underestimate $\phi(M_{BH}, z)$ near the peak of the distribution. However, because low M_{BH} AGN are more likely to be missed by flux-limited surveys, $\phi(M_{BH}, z)$ will be underestimated at low M_{BH} due to incompleteness. The end result is a spurious shift in the inferred BHMF toward higher M_{BH} : incompleteness at low M_{BH} causes one to miss low M_{BH} sources while the intrinsic statistical uncertainty on the broad line mass estimates causes one to overestimate the number of high M_{BH} black holes.
- In § 4.5 we modify the likelihood function to include measurement error in the emission line width. We show that if the measurement errors on the line width are much smaller than the intrinsic physical dispersion in line widths, then measurement error may be neglected. However, if measurement error on the line width is a concern, Equations (56)–(58) should be used for Equation (29) instead of Equation (30).
- We describe in § 5 a Metropolis-Hastings algorithm (MHA) for obtaining random draws from the posterior distribution of the BHMF under the mixture of Gaussian functions model. These random draws may be used to estimate the posterior distribution for the BHMF, as well as to estimate the posterior for any quantities calculated from the BHMF. The posterior provides statistically accurate uncertainties on the BHMF and related quantities, even below the survey detection limits. We use simulation in § 6 to illustrate the effectiveness of our statistical method, as well as to give an example on how to use the MHA output to perform statistical inference.
- We concluded by applying our method to obtain an estimate of the local unobscured quasar BHMF from the $z < 0.5$ BQS quasar sample. Although there is little information in the BQS quasars on the BHMF at $M_{BH} \lesssim 10^8 M_\odot$, the mixture of Gaussian functions estimate suggests that the local quasar BHMF falls off approximately as a power law with slope ~ 2 for $M_{BH} \gtrsim 10^8 M_\odot$ at $z \approx 0.2$. The local quasar BHMF appears to shift toward larger M_{BH} at higher z , and there is marginal evidence for a flattening of the high mass BHMF slope at $z \gtrsim 0.3$. We estimate that at a given M_{BH} , $z < 0.5$ broad line quasars have a typical Eddington ratio of ~ 0.4 and a dispersion in Eddington ratio of $\lesssim 0.5$ dex. However, the estimate of the dispersion in Eddington ratio could be biased toward smaller values if the true distribution of Eddington ratios is significantly skewed toward lower values.

BCK, XF and MV acknowledge support for NSF grants AST 03-07384, 08-06861 and a Packard Fellowship for Science and Engineering. BK acknowledges support by NASA through Hubble Fellowship grant #HF-01220.01 awarded by the Space Telescope Science Institute, which is operated by the Association of Universities for Research in Astronomy, Inc., for NASA, under contract NAS 5-26555. We also acknowledge financial support from HST grants HST-GO-10417 (XF, MV), HST-AR-10691 (MV), and HST-GO-10833 (MV) awarded by the Space Telescope Science Institute, which

is operated by the Association of Universities for Research in Astronomy, Inc., for NASA, under contract NAS5-26555.

REFERENCES

- Adelman-McCarthy, J. K., et al. 2006, *ApJS*, 162, 38
- Andreon, S., Punzi, G., & Grado, A. 2005, *MNRAS*, 360, 727
- Avni, Y., & Bahcall, J. N. 1980, *ApJ*, 235, 694
- Babbedge, T. S. R., et al. 2006, *MNRAS*, 370, 1159
- Barger, A. J., Cowie, L. L., Mushotzky, R. F., Yang, Y., Wang, W.-H., Steffen, A. T., & Capak, P. 2005, *AJ*, 129, 578
- Bentz, M. C., Peterson, B. M., Pogge, R. W., Vestergaard, M., & Onken, C. A. 2006, *ApJ*, 644, 133
- Blanton, M. R., et al. 2003, *ApJ*, 592, 819
- Boroson, T. A., & Green, R. F. 1992, *ApJS*, 80, 109
- Brown, M. J. I., et al. 2006, *ApJ*, 638, 88
- Chib, S., & Greenberg, E. 1995, *Amer. Stat.*, 49, 327
- Collin, S., Kawaguchi, T., Peterson, B. M., & Vestergaard, M. 2006, *A&A*, 456, 75
- Corbett, E. A., et al. 2003, *MNRAS*, 343, 705
- Dempster, A., Laird, N., & Rubin, D. 1977, *J. R. Statist. Soc. B.*, 39, 1
- Di Matteo, T., Springel, V., & Hernquist, L. 2005, *Nature*, 433, 604
- Dunlop, J. S., McLure, R. J., Kukula, M. J., Baum, S. A., O'Dea, C. P., & Hughes, D. H. 2003, *MNRAS*, 340, 1095
- Efron, B., & Petrosian, V. 1992, *ApJ*, 399, 345
- Fan, X., et al. 2001, *AJ*, 121, 54
- Fine, S., et al. 2008, in press at *MNRAS*(arXiv:0807.1155)
- Gebhardt, K., et al. 2000, *ApJ*, 539, L13
- Gelman, A., Carlin, J. B., Stern, H. S., & Rubin, D. B. 2004, *Bayesian Data Analysis* (2nd ed.; Boca Raton:Chapman & Hall/CRC)
- Gelman, A., Meng, X. L., & Stern, H. S. 1998, *Statistica Sinica*, 6, 733
- Gelman, A., Roberts, G., & Gilks, W. 1995, in *Bayesian Statistics* 5, ed. J. M. Bernardo, J. O. Berger, A. P. Dawid, & A. F. M. Smith (Oxford:Oxford University Press), 599
- Greene, J. E., & Ho, L. C. 2007, *ApJ*, 667, 131
- Haehnelt, M. G., & Kauffmann, G. 2000, *MNRAS*, 318, L35
- Hao, L., et al. 2005, *AJ*, 129, 1795
- Hastings, W. K. 1970, *Biometrika*, 57, 97
- Hopkins, P. F., Hernquist, L., Cox, T. J., Di Matteo, T., Robertson, B., & Springel, V. 2006, *ApJS*, 163, 1
- Hopkins, P. F., Richards, G. T., & Hernquist, L. 2007, *ApJ*, 654, 731
- Jester, S., et al. 2005, *AJ*, 130, 873
- Jiang, L., et al. 2006, *AJ*, 131, 2788
- Kaspi, S., Smith, P. S., Netzer, H., Maoz, D., Jannuzi, B. T., & Giveon, U. 2000, *ApJ*, 533, 631
- Kaspi, S., Maoz, D., Netzer, H., Peterson, B. M., Vestergaard, M., & Jannuzi, B. T. 2005, *ApJ*, 629, 61
- Kelly, B. C. 2007, *ApJ*, 665, 1489
- Kelly, B. C., & Bechtold, J. 2007, *ApJS*, 168, 1
- Kelly, B. C., Bechtold, J., Trump, J. R., Vestergaard, M., & Siemiginowska, A. 2008, *ApJS*, 176, 355
- Kelly, B. C., Fan, X., & Vestergaard, M. 2008, *ApJ*, 682, 874 (KFFV08)
- Kollmeier, J. A., et al. 2006, *ApJ*, 648, 128
- Kormendy, J., & Richstone, D. 1995, *ARA&A*, 33, 581
- Krolik, J. H. 2001, *ApJ*, 551, 72
- Kukula, M. J., Dunlop, J. S., McLure, R. J., Miller, L., Percival, W. J., Baum, S. A., & O'Dea, C. P. 2001, *MNRAS*, 326, 1533
- La Franca, F., et al. 2005, *ApJ*, 635, 864
- Little, R. J. A., & Rubin, D. B. 2002, *Statistical Analysis with Missing Data* (2nd ed.; Hoboken:John Wiley & Sons)
- Lynden-Bell, D. 1971, *MNRAS*, 155, 95
- Magorrian, J., et al. 1998, *AJ*, 115, 2285
- Maloney, A., & Petrosian, V. 1999, *ApJ*, 518, 32
- Marconi, A., Axon, D. J., Maiolino, R., Nagao, T., Pastorini, G., Pietrini, P., Robinson, A., & Torricelli, G. 2008, *ApJ*, 678, 693
- Marconi, A., & Hunt, L. K. 2003, *ApJ*, 589, L21
- Marconi, A., Risaliti, G., Gilli, R., Hunt, L. K., Maiolino, R., & Salvati, M. 2004, *MNRAS*, 351, 169
- Marshall, H. L., Tananbaum, H., Avni, Y., & Zamorani, G. 1983, *ApJ*, 269, 35
- Marziani, P., Sulentic, J. W., Zamanov, R., Calvani, M., Dultzin-Hacyan, D., Bachev, R., & Zwitter, T. 2003, *ApJS*, 145, 199
- Matute, I., La Franca, F., Pozzi, F., Gruppioni, C., Lari, C., & Zamorani, G. 2006, *A&A*, 451, 443
- McLeod, K. K., & McLeod, B. A. 2001, *ApJ*, 546, 782
- McLure, R. J., Kukula, M. J., Dunlop, J. S., Baum, S. A., O'Dea, C. P., & Hughes, D. H. 1999, *MNRAS*, 308, 377
- McLure, R. J., & Dunlop, J. S. 2001, *MNRAS*, 327, 199
- McLure, R. J., & Jarvis, M. J. 2002, *MNRAS*, 337, 109
- McLure, R. J., & Dunlop, J. S. 2004, *MNRAS*, 352, 1390
- Merloni, A. 2004, *MNRAS*, 353, 1035
- Merritt, D., & Ferrarese, L. 2001, *ApJ*, 547, 140
- Merritt, D., & Poon, M. Y. 2004, *ApJ*, 606, 788
- Metropolis, N., & Ulam, S. 1949, *J. Amer. Stat. Assoc.*, 44, 335
- Metropolis, N., Rosenbluth, A. W., Rosenbluth, M. N., Teller, A. H., & Teller, E. 1953, *J. Chem. Phys.*, 21, 1087
- Netzer, H. 1990, 20. Saas-Fee Advanced Course of the Swiss Society for Astrophysics and Astronomy: Active galactic nuclei, p. 57 - 160, 57
- Nolan, L. A., Dunlop, J. S., Kukula, M. J., Hughes, D. H., Boroson, T., & Jimenez, R. 2001, *MNRAS*, 323, 308
- Novak, G. S., Faber, S. M., & Dekel, A. 2006, *ApJ*, 637, 96
- Onken, C. A., Ferrarese, L., Merritt, D., Peterson, B. M., Pogge, R. W., Vestergaard, M., & Wandel, A. 2004, *ApJ*, 615, 645
- Page, M. J., & Carrera, F. J. 2000, *MNRAS*, 311, 433
- Percival, W. J., Miller, L., McLure, R. J., & Dunlop, J. S. 2001, *MNRAS*, 322, 843
- Peterson, B. M., et al. 2004, *ApJ*, 613, 682
- Richards, G. T., et al. 2001, *AJ*, 121, 2308
- Richards, G. T., et al. 2006, *AJ*, 131, 2766
- Richstone, D., et al. 1998, *Nature*, 395, A14
- Roeder, K., & Wasserman, L. 1997, *J. Amer. Stat. Assoc.*, 92, 984
- Rubin, D. B. 1981, *J. Educational Statistics*, 6, 377
- Rubin, D. B. 1984, *Annals of Statistics*, 12, 1151
- Schafer, C. M. 2007, *ApJ*, 661, 703
- Schneider, D. P., et al. 2005, *AJ*, 130, 367
- Silk, J., & Rees, M. J. 1998, *A&A*, 331, L1
- Soltan, A. 1982, *MNRAS*, 200, 115
- Spergel, D. N., et al. 2003, *ApJS*, 148, 175
- Steffen, A. T., Barger, A. J., Cowie, L. L., Mushotzky, R. F., & Yang, Y. 2003, *ApJ*, 596, L23
- Schmidt, M. 1968, *ApJ*, 151, 393
- Schmidt, M., & Green, R. F. 1983, *ApJ*, 269, 352
- Shen, Y., Greene, J. E., Strauss, M., Richards, G. T., & Schneider, D. P. 2007, submitted to *ApJ*, (arXiv:0709.3098)
- Tremaine, S., et al. 2002, *ApJ*, 574, 740
- Ueda, Y., Akiyama, M., Ohta, K., & Miyaji, T. 2003, *ApJ*, 598, 886
- Vasudevan, R. V., & Fabian, A. C. 2007, *MNRAS*, 381, 1235
- Vestergaard, M. 2002, *ApJ*, 571, 733
- Vestergaard, M. 2004, *ApJ*, 601, 676
- Vestergaard, M. 2006, *New Astronomy Review*, 50, 817
- Vestergaard, M., & Peterson, B. M. 2006, *ApJ*, 641, 689
- Vestergaard, M., 2008, in progress
- Vestergaard, M., Fan, X., Tremonti, C. A., Osmer, P. S., & Richards, G. T. 2008, *ApJ*, 674, L1
- Waddington, I., Dunlop, J. S., Peacock, J. A., & Windhorst, R. A. 2001, *MNRAS*, 328, 882
- Wandel, A., Peterson, B. M., & Malkan, M. A. 1999, *ApJ*, 526, 579
- Wang, J.-M., Chen, Y.-M., & Zhang, F. 2006, *ApJ*, 647, L17
- Warren, S. J., Hewett, P. C., & Osmer, P. S. 1994, *ApJ*, 421, 412
- Willott, C. J., Rawlings, S., Blundell, K. M., Lacy, M., & Eales, S. A. 2001, *MNRAS*, 322, 536
- Wolf, C., Wisotzki, L., Borch, A., Dye, S., Kleinheinrich, M., & Meisenheimer, K. 2003, *A&A*, 408, 499
- Yu, Q., & Tremaine, S. 2002, *MNRAS*, 335, 965